Understanding Burnout in Healthcare Professionals: Risk Factors for Medical Doctors and Nurses

Leanne Newsham

A thesis submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology

Clinical Psychology Unit, Department of Psychology

The University of Sheffield

December 2025

Declaration

I, the author, confirm this thesis is my own work and that I am aware of the University of Sheffield guidance on unfair means (www.sheffield.ac.uk/new-students/unfairmeans). This thesis is submitted in partial fulfilment of the requirements for the degree of Doctorate in Clinical Psychology. This work has not been submitted for any other degree or to any other institution. No funding has been received for this thesis. No conflicts of interests stated. For any queries about data or code sharing, please contact lnewsham1@sheffield.ac.uk or r.k.webster@sheffield.ac.uk.

Structure and Word Count

Chapter One: Literature Review

Excluding references and tables:	7,999
Including references and tables:	14,076

Chapter Two: Empirical Study

Excluding references and tables:	7,771
Including references and tables:	11,854

Total

Excluding references and tables:	15,770
Including references and tables:	25,930

Lay Summary

Occupational burnout is a psychological response to work-related stress. It is important to understand why healthcare professionals (HCPs) are more susceptible to burnout. Current burnout interventions are beneficial in the short-term, aiding the development of coping skills to manage stressors. However, it is well-recognised that they seldom support the unique needs of individuals. Understanding the risk factors for burnout and exploring how interventions work for different groups is needed to offer effective support. This would help to consider how services can improve their response to burnout to sustain a healthy workforce.

Chapter one provides a systematic review of the literature, exploring whether factors relating to racial-ethnic identity influence how much HCPs experience burnout. Twenty observational studies were identified which examined racial-ethnic related factors and burnout at a single timepoint, of which some also assessed workplace mistreatment. The narrative synthesis indicated that the role of racial-ethnic factors is inconclusive. This aligns with the diverse experiences of underrepresented groups reported in the wider literature. Future research could explore the intersections of individual and sociocultural factors on burnout to provide more supportive and inclusive workplaces to mitigate burnout.

Chapter two reports an empirical study which examined whether different burnout subtypes could predict treatment responses to the Mind Management Skills for Life (MMSFL) programme. Pre-existing data from trials with HCPs in the National Health Service (NHS) identified 12 subtypes using the Oldenburg Burnout Inventory. Results showed subtypes did not influence outcomes; however, subgroups of professionals experienced more severe burnout at the start and end of the intervention. Based on the findings, it is not possible to personalise the MMSFL programme at this stage. Future research should confirm if findings replicate across larger and more diverse samples. Methodological issues and clinical and research implications are discussed for the systematic review and empirical study.

Acknowledgements

Firstly, I would like to thank my supervisors Professor Jaime Delgadillo and Dr Rebecca Webster for your invaluable encouragement and guidance throughout the course of the thesis. You have consistently provided a safe and encouraging learning space and always made time to offer your thoughts, expertise, and support. This has helped me to reach a place where I can say I have done some machine learning! I am grateful to be part of such an important area of research. I would like to extend my gratitude to the researchers and all the healthcare professionals that participated in the clinical trials in the National Health Service (NHS), you made it possible for this thesis to exist.

To my husband and family, a heartfelt thank you for your unwavering support, love, and generosity in life, and especially through the Doctorate in Clinical Psychology journey. Thank you to my parents, Tracy and Paul, and brother Lewis, for always putting a smile on my face and reminding me how proud you are. A special appreciation to my Mum for your words of encouragement and hugs when I needed them most. And especially, thank you to my husband, Nathan, for believing in me and being there every step of the way. You drove miles and miles to allow this to happen and I will be eternally grateful for your love, strength, and patience. I wouldn't be where I am without you.

To my closest friends, thank you for the endless support and loyalty over the years. To my fellow trainees, I am lucky to have been on this journey with you all and couldn't have asked for a kinder and more skilled cohort. We did it! Finally, thank you to the University of Sheffield course team for your compassion through all stages of clinical training.

List of Contents

Chapter One: Literature Review

Racial-Ethnic Related Factors and Occupational Burnout in Nurses and Doctors: A Systematic Review and Narrative Synthesis

Abstract

Objective: Burnout among healthcare workforces is a well-recognised problem. However, an understanding of the role of racial-ethnic identity in burnout among registered doctors and nurses is not readily available. This review aimed to identify whether racial-ethnic related factors are associated with burnout, and to consider research and clinical implications.

Methods: A systematic review (PROSPERO ID: CRD420251110070) was conducted through searching six databases (CINAHL, Medline, PsycINFO, PsycARTICLES, Scopus, Web of Science) in September 2025. The Population, Exposure, and Outcome framework guided eligibility criteria; peer-reviewed, English language articles describing quantitative research that assessed burnout prevalence among doctors and nurses and tested associations with at least one racial-ethnic related predictor. Study quality was assessed using The Mixed Methods Appraisal Tool. Predictors of outcome were summarised using a narrative synthesis.

Results: Twenty eligible observational studies investigated racial-ethnic group as a predictor of burnout, of which five also examined experiences of mistreatment. Largely, the Maslach Burnout Inventory was used to measure burnout. The relationship between burnout and racial-ethnic related factors is mixed; however, nuanced group-specific trends were found.

Conclusions: This review highlights the need for additional research focusing on burnout conceptualisation, measurement, and prevention, considering the intersection of individual, sociocultural, and organisational factors in burnout experiences.

Keywords: *Burnout, race, ethnicity, mistreatment, healthcare professionals, predictors*

Practitioner Points

- The role of racial-ethnic related factors in burnout is inconclusive.
- Future burnout research is needed to examine the intersection of sociocultural factors.
- Inclusive responses to burnout should be prioritised in healthcare organisations to address disparities among underrepresented groups.

Introduction

Prevalence and Impact of Healthcare Burnout

Occupational burnout is a psychological response to prolonged work-related stressors, characterised by physical, emotional, and mental exhaustion (World Health Organization, 2019). Burnout is a global phenomenon, with healthcare professionals (HCPs) experiencing higher rates of burnout compared to other occupations, mainly due to immense workloads, rising clinical complexity, and workforce shortages in already under-resourced systems. According to the latest National Health Service (NHS) workforce survey (NHS Employers, 2025), 41% of staff felt unwell as a result of their work, with 56% of staff reporting presenteeism (attending work while ill). The severe impact of burnout is widespread, leading to higher rates of depression, anxiety, and suicidal ideation among HCPs (Salvagioni et al., 2017; West et al., 2018), as well as decreased patient satisfaction, poorer quality of care, and increased risk of medical errors (Delgadillo et al., 2018; Hall et al., 2016; Li et al., 2024). At an organisational level, burnout drives absenteeism and high staff turnover, creating a vicious cycle of unrelenting pressure and greater workload demands for existing staff, only perpetuating workforce and financial challenges in healthcare services (Nash, 2025).

The COVID-19 pandemic marked a pivotal point of soaring demands and service disruption, placing unprecedented pressures on HCPs (Harding et al., 2025). This drove burnout within the NHS workforce to critical levels, exacerbating pre-existing issues. Medical and nursing professionals are particularly at risk due to their high workloads and the emotional strain of clinical work. Globally, the pooled-prevalence of burnout among nurses has been estimated at 11% (Woo et al., 2020). Recent figures suggest a third of nurses experienced emotional exhaustion during the pandemic (Getie et al., 2025). Medical doctors consistently report some of the highest burnout levels with similar trends observed in studies from Europe, the Middle East and North Africa, and South America (Karuna et al., 2022;

Macaron et al., 2023; Rotenstein et al., 2018). Comparable rates of burnout have been noted among doctors and nurses in the United Kingdom (UK; Kinman et al., 2023). National NHS surveys report 40% of nurses as burned out, alongside a third of doctors feeling unable to manage their weekly workload (NHS Employers, 2025). Critically, 61% of doctors in training are considered at high risk of burnout (General Medical Council [GMC], 2025).

Racial-Ethnic Diversity in Healthcare Workforces

Race is defined as the social construction of people based on perceived physical traits and ethnicity represents shared cultural characteristics (e.g., religion, language, values). While race and ethnicity are often used interchangeably in research, the American Psychological Association (APA, 2019a; 2019b) advise racial-ethnic identity when referring to a range of groups to ensure bias-free language, along with underrepresented group (URG) when describing a subgroup with different sociocultural characteristics than the majority population. A quarter of NHS staff self-identify from URGs (e.g., Asian, Black, multiracial), making the workforce more diverse than at any point in its history (NHS England, 2023). Asian NHS doctors form the largest URG, comprising a third of junior and consultant roles (Stockton & Warner, 2024). Underrepresented nurses make up 39% of the NHS workforce with continued growth anticipated (Rolewicz et al., 2024). National UK training surveys report higher burnout rates in medical trainees identifying from URGs, particularly in emergency and renal medicine (GMC, 2025; Graham-Brown et al., 2021).

Workforce trends in the United States (US) indicate that nurses and doctors from URGs (e.g., Asian, Black, Hispanic) remain underrepresented relative to national and local populations despite increasing recruitment rates over the past decade (Hynson et al., 2022; Kim et al., 2025; Smiley et al., 2025). Other well-represented surveys in the US show greater job dissatisfaction and intent to leave among Black nurses (Carthon et al., 2021), whereas underrepresented doctors report less burnout than majority groups (Odei & Chino, 2021). A systematic review found inconsistent burnout differences among underrepresented medical professionals (i.e., medical academic staff, doctors, students) in North America, showing greater rates of burnout in some studies and lower prevalence or no differences in others (Lawrence et al., 2021). Based on these inconsistencies and wider literature suggesting that burnout might be expressed differently across diverse groups (Bafna et al., 2025), it is critical to explore the drivers of burnout across different racial-ethnic identities.

Underrepresented HCPs often face unique and negative workplace challenges, such as discrimination, microaggressions, and a lack of institutional support (Mosley et al., 2025). The Equality Act (2010) defines racial discrimination as a direct or indirect act that treats an individual differently based on their racial-ethnic identity, such as institutional racism (discrimination embedded in policies). Incidents of discrimination towards underrepresented nurses have doubled over the past three years in the UK (Royal College of Nursing [RCN], 2025) and have been linked to perceived obstacles to success in medical students, as well as burnout and suicidal thoughts among US doctors (Hu et al., 2019; Odom et al., 2007). Higher rates of harassment, bullying, and abuse are widely reported among URGs in healthcare settings; however, Black, Asian, and multiracial doctors often underreport experiences of racism due to a lack of confidence in reporting procedures and fear of negative repercussions (e.g., career progression). This suggests prevalence rates are likely higher (British Medical Association [BMA], 2022; GMC, 2025).

Evidently, healthcare workforces are becoming increasingly more diverse, requiring an understanding of burnout in the context of racial-ethnic identities. Underrepresented HCPs are more likely to remain in entry-level jobs or work in under-resourced environments, exacerbating existing workplace pressures (Wilbur et al., 2020). Yet, diverse workplaces where cultural preferences are supported may be protective against burnout, highlighting the complex relationship between individual characteristics and burnout (Bafna et al., 2025;

Syahir et al., 2025). Burnout undoubtedly has implications for work satisfaction, absenteeism, and staff retention, emphasising the critical need to examine burnout across racial-ethnic identities in order to understand its impacts and identify points for prevention and intervention. Despite a broad and widely documented awareness of burnout among HCPs, there remains limited research on the experiences of individual racial-ethnic groups, particularly among high-risk professionals such as doctors and nurses.

Burnout Conceptualisation

Burnout was first identified in the 1970s by Herbert Freudenberger, defined as physical and emotional exhaustion (Freudenberger, 1975). Since, the Maslach Burnout Inventory (MBI) has provided a standardised way to conceptualise and measure workplace burnout (Maslach & Jackson, 1981). The MBI remains the most well-established measure, comprising three dimensions of emotional exhaustion (depletion of emotional resources), depersonalisation (cynical attitudes), and personal accomplishment (feeling ineffective). The original 22-item MBI-Human Services Survey (MBI-HSS) was designed for health, social work, and education professionals, with later versions modified for healthcare professionals (MB-HSS-Medical Personnel). Both consistently demonstrate high validity and reliability (García et al., 2019; Lin et al., 2022; Poghosyan et al., 2009) across several occupations and countries (e.g., Canada, UK, US). High levels of emotional exhaustion and depersonalisation are commonly used to define burnout in the literature as some studies show weak correlations between personal accomplishment and burnout outcomes (West et al., 2018).

Additional tools have been developed to address limitations with the MBI-HSS, mainly translation inaccuracies (e.g., loss of original meaning of positive/negative items) that have led to challenges with cross-cultural utility (Shoman et al., 2021; Soares et al., 2023). The Oldenburg Burnout Inventory (OLBI) was developed to examine burnout as a continuum, using two domains (exhaustion and disengagement) with positive and negatively framed items (Demerouti et al., 2001). The Copenhagen Burnout Inventory (CBI; Kristensen et al., 2005) conceptualises burnout as fatigue and exhaustion, intending to clarify how burnout is defined. However, minimising burnout to a singular construct of exhaustion has been criticised as neglecting other key aspects (Reis et al., 2015). While the MBI, OLBI, and CBI are most commonly used, other measures such as the Professional Quality of Life (ProQol; Stamm, 2010), single-item, and researcher-developed tools are employed in burnout research. The clinical utility of these measures has not been examined among URGs.

Current Study

Burnout has been declared a public health issue by the World Health Organization (2019) with adverse impacts remaining a significant concern for health organisations that continue to face workforce strains exacerbated by the COVID-19 pandemic. While research has highlighted disparities in burnout across demographic and occupational contexts, racial-ethnic related factors are inconsistently examined in the literature, leaving a gap in consolidated knowledge. Emerging evidence indicates that racial-ethnic identity and workplace mistreatment may influence burnout risk; however, a comprehensive synthesis of the literature has not focused on high-risk professionals such as registered doctors and nurses. The main aim of this systematic review was to synthesise quantitative evidence on burnout prevalence and its associations among doctors and nurses from a range of racial-ethnic identities. The primary objective was to identify key racial-ethnic related factors contributing to burnout, considering their implications for research and clinical practice. Where possible, results have been disaggregated by specific racial-ethnic identities.

Methods

Study Protocol

This systematic review has been reported in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al., 2021). The PRISMA reporting guideline can be reviewed in Appendix A. The review protocol was registered and published in the International Prospective Register of Systematic Reviews (PROSPERO) prior to conducting database searches (PROSPERO 2025: CRD420251110070). No funding was received for this review and the authors report no conflicts of interest.

Eligibility Criteria

Using the Population, Exposure, and Outcome framework (PEO; Moola et al., 2015), studies were assessed against several eligibility criteria for inclusion (see Table 1). The PEO framework ensured comprehensive coverage of studies examining whether naturally occurring exposures (racial-ethnic factors) are associated with burnout (outcome) in medical doctors and nurses (population). This review will utilise racial-ethnic identity when referring to a range of groups (e.g., White, Asian, Black) and underrepresented group (URG) when describing groups other than the majority population. To remain inclusive of all racial-ethnic identities, commonly accepted designations (e.g., census categories) will be used while being sensitive to the preferred designations reported by authors (APA, 2019b).

Table 1

Inclusion and Exclusion Criteria

	Inclusion	Exclusion
Population	Registered medical doctors and nurses working in patient or clinical settings, including all specialties and job grading.	Non-healthcare professionals (e.g., admin, clerical, students) or allied health professionals.
Exposure(s)	Racial-ethnic related factors, including self-identified race or ethnicity, religion, and experiences of mistreatment in relation to racial-ethnic identity.	Exposures that do not relate to racial-ethnic identity (e.g., gender-based discrimination).
Outcomes	Validated measure of burnout symptoms (e.g., MBI, CBI, OLBI) administered at least once.	Outcomes solely measuring psychological disorders (e.g., depression) or work-related performance without a direct assessment of burnout.

		Burnout symptoms not measured using a validated measure.
Setting	Any outpatient or inpatient setting.	Non-healthcare or academic setting.
Study Design	Quantitative research that assessed prevalence of burnout and tested associations.	Qualitative research, grey literature, conference proceedings, presentations or media articles.
	Experimental and non-experimental study designs.	Publications not written in the English language. Non-accessible studies.

Note. MBI = Maslach Burnout Inventory, CBI = Copenhagen Burnout Inventory, OLBI = Oldenburg Burnout Inventory.

Search Strategy

In line with the core methodological steps outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2024), a search strategy was developed. Six databases (CINAHL, Medline, PsycINFO, PsycARTICLES, Scopus, Web of Science) were searched on 5th September 2025 to identify relevant articles, using all of the search terms listed in Appendix B. Database selection was guided by the review topic to optimise coverage of psychological, medical, and nursing literature. Key search terms relating to racial-ethnic identity, healthcare roles, burnout, and study design were combined with a Boolean operator (OR) to maximise the retrieval of pertinent articles. To increase the scope of the search, the review aimed to include studies conducted globally and no date limits were applied. The search strategy did not require adjustment and a subsequent search was not undertaken as the initial search date was within six months of submission, ensuring inclusion of the most up to date evidence.

Selection of Articles

The reference management software EndNote was used to perform the study selection process (Bramer et al., 2017). Database search results were imported and merged. Duplicates were detected using the duplicate detection function, alongside manual verification by the primary reviewer prior to removal to ensure accuracy. A two-stage screening process involved initial title and abstract screening, followed by full-text reviews. The primary reviewer screened all titles and abstracts for eligibility. A second reviewer independently screened 20% of randomly selected titles and abstracts to evaluate inter-rater agreement; there was full (100%) agreement. A full-text review was completed for remaining articles in accordance with the predefined eligibility criteria by the primary reviewer. Forward and backward citation searching was carried out for all included studies to identify any additional articles not captured in the database search. The study selection process is outlined in Figure 1 using a PRISMA flow diagram (Page et al., 2021).

Data Extraction

Data from the included studies were extracted in line with the Cochrane Collaboration Data Collection guidance (Higgins et al., 2024), using a table that was developed and tailored for the specific aim and objectives of this review. This ensured a structured approach to the recording of study characteristics and outcomes. Data was extracted and tabulated by the primary reviewer, including citation, study design, country of study, main characteristics of participants (i.e., sample size, mean age/frequency, gender, racial-ethnic identity), population (i.e., HCPs, speciality), study setting, outcome measure, variables assessed as predictors (i.e., race, ethnicity, experiences of mistreatment), and results (i.e., factors associated with burnout). Results from individual studies are presented in tables for each predictor category along with effect sizes and available significance values. No eligible studies were identified that examined religion as a predictor of outcome.

Quality Assessment

Risk of bias for eligible studies was assessed using the Mixed Methods Appraisal Tool (MMAT; Hong et al., 2018). The MMAT (see Appendix C) is a validated tool designed to assess the quality of empirical studies across a range of study designs. This review utilised the quantitative non-randomised category. Feasibility of appraisal is determined with two screening questions, followed by five design-specific questions. The MMAT does not provide a definitive risk of bias classification, rather each individual criterion is rated as "yes" (fully met the criterion), "no" (did not meet the criterion), or "can't tell" (insufficient information). In line with Webster et al. (2019), a modified scoring approach facilitated narrative comparison of study quality. The original "yes", "can't tell", and "no" answers correspond to "low", "unclear", or "high" risk of bias, respectively. The MMAT strongly discourages the calculation of a composite score from the ratings of each criterion. Rather, a detailed overview of each criterion is advised to better inform judgements about the quality of the evidence. A second reviewer independently assessed 20% of the included studies. Inter-rater reliability was assessed using Cohen's Kappa (Cohen, 1960), indicating substantial agreement, k = .73. Any discrepancies were discussed and resolved.

Data Synthesis

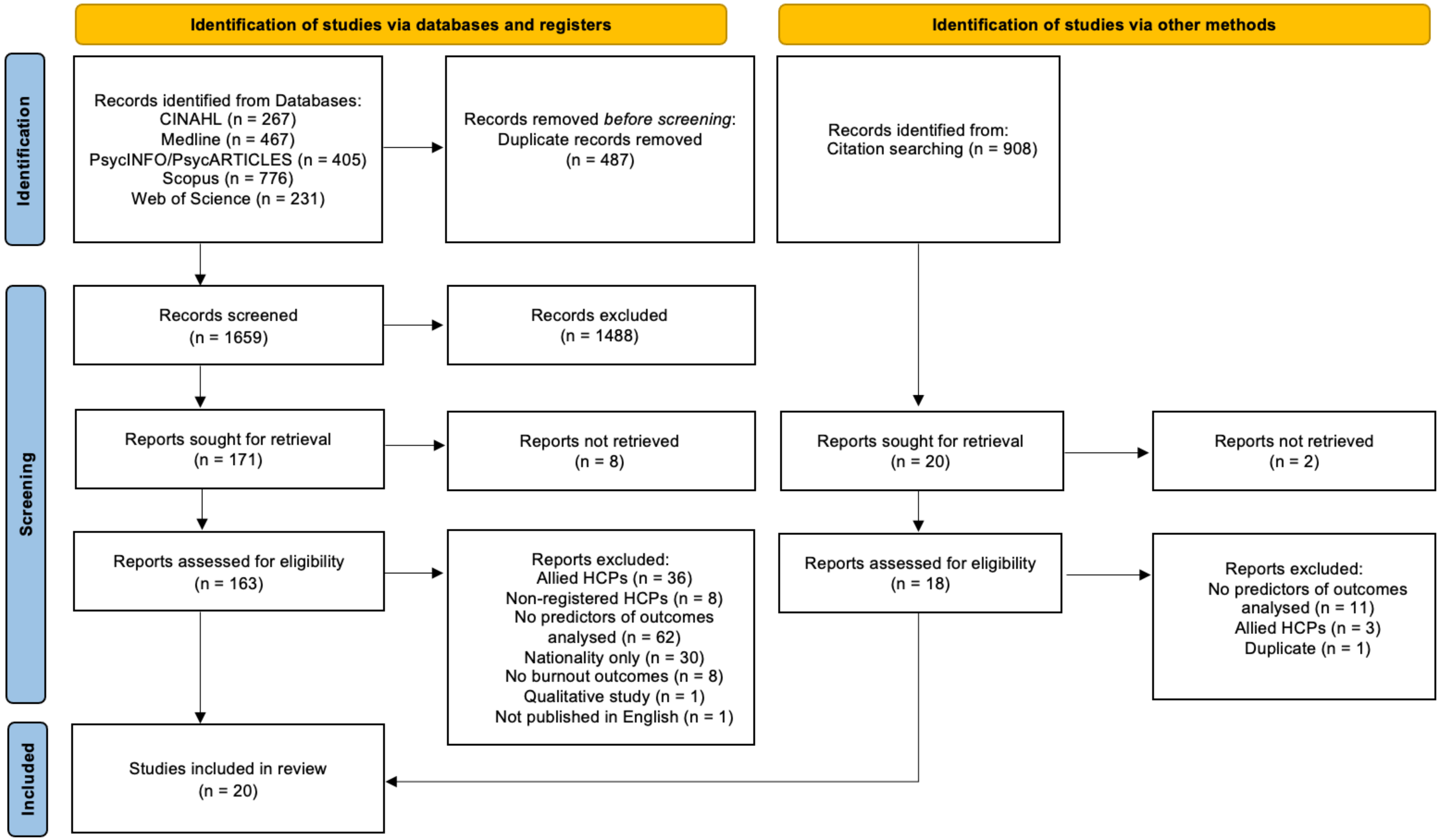
Narrative synthesis was the method used for summarising and explaining the key study findings. Meta-analytic methods could not be applied based on the heterogeneity of study designs, measurement tools, and statistical reporting of outcomes, meaning a pooled analysis was not appropriate (Martinez et al., 2025). Included studies were all observational and varied in how racial-ethnic identity were defined and analysed as variables. Additionally, burnout was measured using different or adapted versions of the same tool (e.g., MBI) with a range of effect size measures across studies. There is no agreed consensus for conducting a narrative synthesis; however, Popay et al. (2006) provide four main elements that have proven influential in health research. These include theory development, creating a preliminary synthesis, exploring relationships within and between studies, and assessing the robustness of the synthesis. Accordingly, a research question was defined; studies were organised based on predictors, professional group, and outcome measure; connections within and between studies were explored; and the robustness of findings was evaluated using a weight of evidence approach, considering trends and study quality.

Results

The search strategy generated 2146 titles and abstracts, of which 487 were removed as duplicates. The main reasons for study exclusion during initial title and abstract screening were non-healthcare populations, no evidence of burnout outcomes, no racial-ethnic related predictors, or qualitative methodology. This identified 163 articles for full-text review. Of these, 146 articles were excluded based on not meeting eligibility criteria. Common reasons included mixed populations of non-registered (e.g., students, nursing assistants, admin staff) or allied HCPs, no predictors of outcomes were analysed, article was not available in English, or only nationality was listed. Seventeen articles were eligible for inclusion in the review. Forward and backward citation searching of included studies produced an additional 908 articles, of which 3 met the eligibility criteria following full-text screening. As a result, 20 studies were included in the review. Figure 1 presents a PRISMA diagram summarising the screening process and reasons for exclusion.

Figure 1

PRISMA Diagram for Selection of Articles



Study Characteristics

There was a total of (k = 20) observational studies in the review, including (k = 19) cross-sectional and (k = 1) cohort. Studies were published between 2019 and 2025, and conducted across several countries, including Brazil (k = 1), Canada (k = 2), United Kingdom (k = 3), and the United States of America (k = 14). Sample sizes among studies ranged from (n = 104) to (n = 7,440), with a total pooled sample size of (N = 40,084). The general age range of participants was 30 to 75 years old. Two studies did not report age, and one did not report gender. Across the 20 studies, racial and ethnic demographics were reported for 35,711 participants, with around 70% self-identifying as White and 30% from URGs (e.g., Asian, Black, Hispanic, Other). The proportion of White participants ranged from 29% to 81% across individual studies. Several studies grouped participants as "underrepresented" (e.g., American Indian, Black, Hispanic/Latinx, multiracial), whereas others separated Asian and South Asian into subgroups or used broader classifications such as non-White and non-Black, indicating substantial differences in sample diversity.

The majority of studies (k = 14) focused on medical doctors from a range of specialities (e.g., Emergency Medicine, Obstetrics and Gynaecology, Orthopaedic) and training grades. In UK studies, consultant referred to senior doctors who have completed specialist training (BMA, 2024), equivalent to an attending physician in the US. Across studies, resident doctor was used as a broad term for doctors in postgraduate training, with registrars and fellows representing more advanced stages. In this review, consultant, attending physician, specialist doctor, resident, registrar, and fellow were collapsed into a single category (medical doctor), as all titles represent fully registered (qualified) doctors. No subgroup analyses of training grade were completed. Other studies focused on nurses (k = 4), of which one consisted of mental health nurses. Remaining studies (k = 2) had a mixed sample of medical doctors and nurses. Hospital settings (e.g., acute care, maternity) were the most common workplace. However, four studies reported community-based (e.g., primary care) contexts, and eight studies did not specify the study setting.

All studies included a validated burnout measure. The MBI-HSS was employed in (*k* = 17) studies, of which eight used the 22-item standard measure, five the medical personnel version (MBI-HSS-MP), two the 12-item abbreviated version (aMBI), and two used a rapid screening tool. Three remaining studies employed either the Mini Z single-item instrument, the ProQol, or the CBI. In addition, five studies used self-report mistreatment measures, including the Racial Microaggressions Scale (RMAS) and researcher-developed self-report questionnaires (e.g., types and sources of mistreatment). The most common statistical method used to examine predictors was multivariable logistic regression (*k* = 9), followed by Poisson regression (*k* = 2) and multiple linear regression (*k* = 2). Other studies used a range of statistical methods, including ordinal and multilevel logistic regression, chi-square tests, multilevel ANOVA, and hierarchical multiple regression. See Table 2 for full study characteristics.

Table 2

Summary of Study Characteristics for Included Articles

Authors and Year	Design	Country	Sample Size (N)	Age Mean (SD) / Frequency	Gender	Racial-Ethnic Group N (%)	Population (Speciality)	Setting	Outcome Measures
Agha et al. (2024)	Cross-Sectional	UK	104	<31 = 8 31-35 = 60 36-40 = 25 41-45 = 9 >45 = 2	F = 46	White = 43 (41.4) Asian = 49 (47.1) Black = 1 (1.0) Arab = 3 (2.9) Other = 4 (3.9) Unknown = 4 (3.9)	Specialist Registrars (Diabetes and Endocrinology)	-	Burnout (MBI-HSS) Self-Report Mistreatment Questionnaire
Akinleye et al. (2024)	Cohort	US	454	≤45 = 195 46-55 = 134 ≥56 = 123	F = 303 M = 127	Non-Hispanic White = 291 (32.6) Asian = 48 (22.2) Black, non-Hispanic = 33 (32.4) Hispanic = 27 (34.2) Other = 26 (28.9) Unknown = 29 (24.4)	Physician Nurse Practitioner	Primary Care	Burnout (Mini Z)
Bourne et al. (2019)	Cross-Sectional	UK	3073	43.3 (13.6)	F = 2069 M = 963	White = 1767 (57.9) Asian = 832 (27.3) Black = 201 (6.6) Multiracial = 172 (5.6) Other = 82 (2.7)	Consultants Trainees Specialist Doctors (Obstetrics and Gynaecology)	-	Burnout (MBI-HSS-MP)
Carthon et al. (2024)	Cross-Sectional	US	7440	42.5 (13.2)	F = 6713	Non-Hispanic White = 6642 (89.3) Hispanic = 798 (10.7)	Nurses	Acute Care Hospitals	Burnout (MBI-HSS; EE subscale)
Coombs et al. (2020)	Cross-Sectional	US	146	30	F = 61 M = 84	White = 116 (80.6) Asian = 5 (3.5) Indian/South Asian = 4 (2.8) Black/African = 2 (1.4) Latinx = 2 (1.4) Hispanic = 7 (4.9) Other = 8 (5.6)	Resident Physicians (Plastic and Reconstructive Surgery)	-	Burnout (MBI-HSS)

Douglas et al. (2021)	Cross-Sectional	US	3096	43.4 (10.5)	-	Underrepresented: Hispanic = 231 (51.3) Black = 204 (45.3) American Indian = 30 (6.7)	Physicians (Obstetrics)	Inpatient	Burnout (MBI-HSS; EE, DP)
Dyrbye et al. (2022)	Cross-Sectional	US	6512	<35 = 225 35-44 = 1363 45-54 = 1647 55-64 = 1856 ≥65 = 1203	F = 2450 M = 4062	Non-Hispanic White = 3622 (70.5) Black, non-Hispanic = 181 (3.5) Non-Hispanic AAPI = 681 (13.3) Indigenous = 182 (3.5) Non-Hispanic Other = 102 (2.0) Hispanic/Latinx = 369 (7.2)	Physicians	Hospital Community Private	Burnout (MBI-HSS; EE, DP) Six-Item Mistreatment Questionnaire
Garcia et al. (2020)	Cross-Sectional	US	4424	52.46 (12.03)	F = 1688 M = 2722 Other = 5	Non-Hispanic White = 3480 (78.7) Hispanic/Latinx = 278 (6.3) Black, non-Hispanic = 124 (2.8) Asian, non-Hispanic = 542 (12.3)	Physicians (Multiple Specialities)	-	Burnout (MBI-HISS; EE, DP)
Khan et al. (2021)	Cross-Sectional	Canada	249	25-35 = 32 36-50 = 124 51-65 = 72 ≥66 = 19	F = 122	White = 166 (68.0) Asian/Pacific Islander = 41 (16.8) South Asian = 19 (7.8) Other = 18 (7.4)	Physicians	Hospitals	Burnout (MBI-HSS-MP)
Lee et al. (2021)	Cross-Sectional	US	335	49 (11.0)	F = 21 M = 79	White = 234 (78.2) Black = 14 (4.7) Native/Indian American = 6 (2.0) Asian/Pacific Islander = 27 (9.0) Other = 18 (6.0)	Residents Fellows Surgeons (Orthopaedic Surgery)	-	Burnout (MBI-HSS-MP)
Lu et al. (2023)	Cross-Sectional	US	7680	-	F = 2698 M = 4768	Non-Hispanic White = 4919 (64.0) Underrepresented: Black = 282 (3.7) Asian/Pacific Islander = 1042 (13.6) Hispanic/Latinx = 329 (4.3) Native American/Alaska Native = 33 (0.4) Other/Multiracial = 934 (12.2) Unknown = 141 (1.8)	Residents (Emergency Medicine)	-	Burnout (MBI-HSS; 2-item EE, DP) Self-Report Mistreatment Questionnaire
Martinez et al. (2022)	Cross-Sectional	US	1723	<40 = 179 40-49 = 579	F = 599 M = 1124	White = 1240 (72.0) Non-White = 365 (21.2)	Physicians	-	Burnout (MBI-HSS)

				50-59 = 480 ≥60 = 485					
Merces et al. (2020)	Cross-Sectional	Brazil	1125	≤35 = 587 ≥36 = 538	F = 989 M = 136	Black = 852 (77.6) Non-Black = 246 (22.4)	Nurses	Primary Care	Burnout (MBI-HSS)
Murray and Chiotu (2024)	Cross-Sectional	Canada	148	39.74 (11.8)	F = 97 M = 35 Other = 3	White = 47 (34.8) Asian = 51 (37.8) Black = 24 (17.8) Indigenous = 13 (9.6)	Mental Health Nurses	Hospital	Burnout (ProQol)
Nituica et al. (2021)	Cross-Sectional	US	682	<35 = 601 ≥35 = 81	F = 383 M = 299	White = 458 (67.2) Asian/Pacific Islander = 113 (16.6) Hispanic = 47 (6.9) Multiracial/Other = 37 (5.4) Black = 27 (4) American Indian/Alaskan Native = 1 (<1)	Resident Physicians	Hospital	Burnout (aMBI)
Parker et al. (2025)	Cross-Sectional	UK	1114	<35 = 306 35-49 = 479 ≥50 = 329	F = 866 M = 229	White = 709 (63.6) Asian = 251 (22.5) Black = 58 (5.2) Multiracial = 24 (2.2) Other = 50 (4.5)	Consultants Specialist Doctors Trainees (Obstetrics and Gynaecology)	-	Burnout (MBI-HSS-MP)
Pillado et al. (2022)	Cross-Sectional	US	510	-	M = 329	White = 223 (53.1) Asian = 102 (24.4) Hispanic/Latinx = 33 (7.6) Black = 16 (4.2) Other = 41 (10.8)	Physicians (Vascular Surgery Trainees)	-	Burnout (aMBI) Six-Item Racial Discrimination Questionnaire
Robinson et al. (2024)	Cross-Sectional	US	150	21-34 = 70 35-54 = 65 55-64 = 12 65-75 = 3	F = 139 M = 11	White = 96 (62.7) Black = 34 (22.7) Asian/Pacific Islander = 11 (7.3) Hispanic/Other = 11 (7.3)	Attending Physicians Resident Physicians Nurses (Obstetrics and Family Medicine)	Hospital (Maternity Units)	Burnout (MBI-HSS-MP)

Sudol et al. (2021)	Cross-Sectional	US	588	20-39 = 150 40-49 = 224 50-59 = 150 ≥60 = 64	F = 259 M = 329	White = 221 (37.6) Hispanic = 31 (5.3) Asian = 189 (32.1) South Asian = 47 (8.0) Middle Eastern = 39 (6.6) Black = 23 (3.9) Pacific Islander = 4 (0.7) Multiracial = 34 (5.8)	Surgeons and Anaesthesiologists (Surgery)	Hospital	Burnout (MBI-HSS) Racial Discrimination (RMAS)
Taylor et al. (2024)	Cross-Sectional	US	531	-	F = 368 M = 148	White = 155 (29.2) Black = 52 (9.8) Asian = 137 (25.8) Hispanic/Latinx = 131 (24.7) Native American = 6 (1.1) Multiracial = 31 (5.8) Other = 19 (3.6)	Nurses	Acute Care Hospitals	Burnout (CBI; Work-Related Subscale)

Note. UK = United Kingdom; US = United States; MBI-HSS = Maslach Burnout Inventory-Human Services Survey; MBI-HSS (MP) = Maslach Burnout Inventory-Human Services Survey for Medical Personnel; aMBI = Abbreviated Maslach Burnout Inventory; CBI = Copenhagen Burnout Inventory; ProQol = Professional Quality of Life Scale; RMAS = Racial Microaggression Scale (RMAS); EE = Emotional Exhaustion (MBI-HSS, 9-items); DP = Depersonalisation (MBI-HSS, 5-items); F = Female; M = Male; SD = Standard Deviation; Non-Hispanic AAPI = Non-Hispanic Asian American and Pacific Islander.

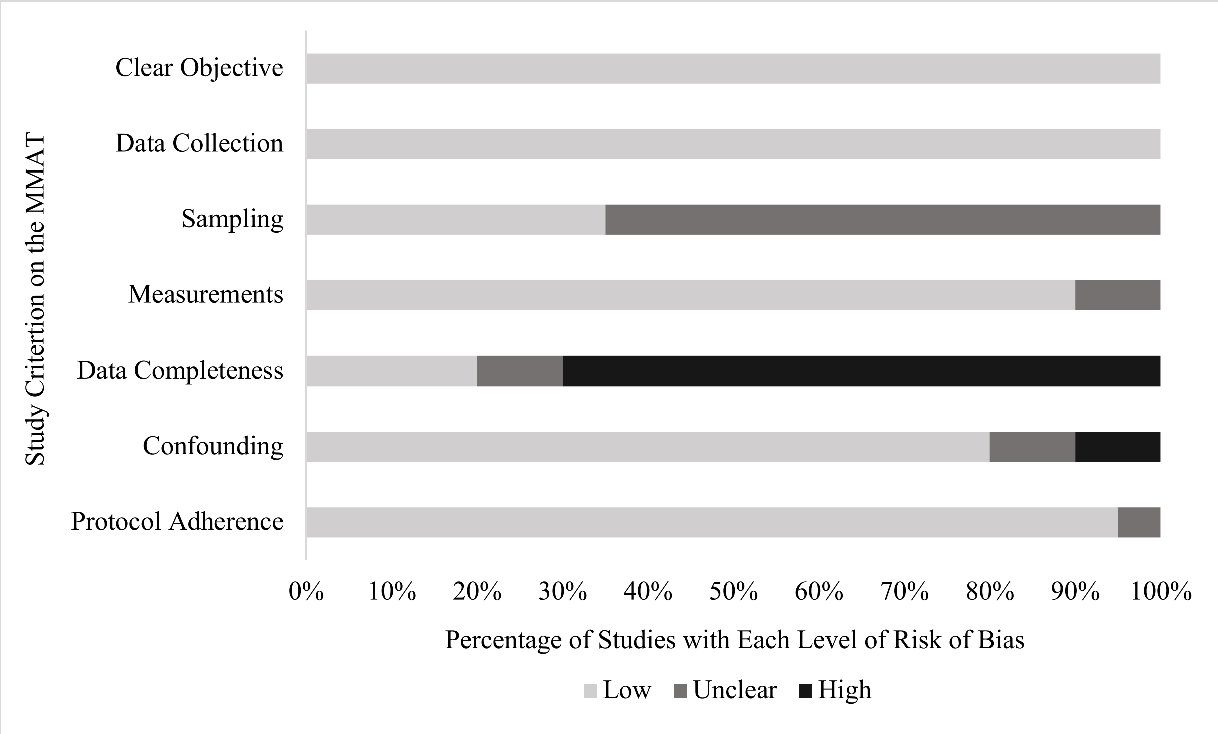
Risk of Bias Assessments

In line with the MMAT guidance (Hong et al., 2018), the criterion-by-criterion ratings were used to judge the quality of the evidence. Overall, the quality of the 20 observational studies was low to moderate (see Figure 2). Clear objectives were reported in all studies, and the data collected was appropriate in addressing the objectives. However, sampling methods in the majority of studies were inadequate and vulnerable to selection bias as participants were often selected based on their availability around particular criteria (e.g., healthcare profession, setting, speciality). Low response rates (below 50%), small sample sizes, or the absence of disaggregated categories for racial and ethnic groups reduced the representativeness of studies. Additionally, a large number of studies did not compare sample data (i.e., percentage of racial-ethnic groups) to the known proportions in the target population. For example, several studies report the number of underrepresented participants as lower or higher than local populations without describing methods to achieve this, whereas others have compared sample demographics with reliable sources (e.g., professional organisations, surveys) to understand the distribution of certain characteristics.

Validated and reliable burnout outcome measures were used in the majority of studies; however, these self-administered tools raise questions regarding social desirability bias. Similarly, mistreatment measures were largely researcher-developed and prone to measurement error as they were not validated against a reliable standard for data collection within the target populations, contributing to further issues with validity and generalisability. Largely, the studies (80%) accounted for confounding variables with appropriate statistical methods, increasing the accuracy of outcomes. Conversely, only four studies reported complete outcome data with the majority of other studies excluding missing values from analyses, and as such were at heightened risk of biased estimates. Two studies used a valid imputation method for handling the missing data. Overall, all studies were administered as intended and while some considered the impact of the COVID-19 pandemic, there was one study that had not anticipated the influence on outcomes. A summary of the risk of bias assessments can be found in Appendix D.

Figure 2

Quality of Included Observational Studies



Note. MMAT = Mixed Methods Appraisal Tool

Predictors of Outcome

Racial-ethnic identity was examined as a predictor of burnout outcomes in all studies. Five of these studies (Agha et al., 2024; Dyrbye et al., 2022; Lu et al., 2023; Pillado et al., 2022; Sudol et al., 2021) also explored the associations among racial-ethnic identity, experiences of mistreatment, and burnout. Religion was not examined as a predictor in any of the studies. The categorisation of racial-ethnic identity, burnout measurement tools, and statistical methods varied considerably across studies, meaning a robust meta-analysis could not be conducted. As a result, a narrative synthesis of findings was undertaken in line with the published protocol for this review. Findings from included studies are grouped based on the nature of predictors, healthcare professional role, and outcome measures. Five studies are repeated within the results tables as their predictors were examined on more than one outcome measure. For example, Sudol et al. (2021) assessed predictors against the MBI-HSS and the RMAS; however, the same participants were used.

Racial-Ethnic Identity and Burnout

Twenty studies examined racial-ethnic identity as a predictor of burnout outcomes in samples of medical doctors or nurses. Two of these studies used a combined sample of doctors and nurses (Akinleye et al., 2024; Robinson et al., 2024). Effect sizes, available p-values, and interpretation of results (i.e., direction of effect) for each of the predictors are summarised in Table 3. The majority of studies used White or non-Hispanic White as the reference group, and several used broader classifications to describe racial identity (e.g., non-Black, non-White), limiting the generalisability of results for specific racial-ethnic identities. Fourteen studies reported statistically significant associations between racial-ethnic identity and burnout.

Studies Comparing Burnout Outcomes with White Reference Groups

Thirteen studies examined racial-ethnic identity as a predictor of burnout, using White or non-Hispanic White as the reference group. These studies comprised samples of doctors or nurses, apart from one which used a combined sample (Akinleye et al., 2024). Across studies, Asian participants frequently showed lower odds of burnout, with effect sizes ranging from small to moderate (Akinleye et al., 2024; Bourne et al., 2019; Dyrbye et al., 2022; Garcia et al., 2020). In two of these studies, a similar trend was found for the Hispanic/Latinx group who were less likely to report burnout (Dyrbye et al., 2022; Garcia et al., 2020). Garcia et al. (2020) also reported lower odds of burnout among Black participants, demonstrating the strongest effect size across URGs in the study. Lu et al. (2023) collapsed racial-ethnic identities into a single underrepresented category revealing a small protective effect for burnout risk. Similarly, Nituica et al. (2021) showed lower emotional exhaustion or depersonalisation among Asian, Black, and Hispanic groups.

Two studies classified underrepresented participants as non-White, reporting lower overall burnout (Lee et al., 2021; Martinez et al., 2022). In addition, Martinez et al. (2022) examined burnout differences in two separate cohorts (2013-2014; 2019-2020), demonstrating a lower likelihood of emotional exhaustion and depersonalisation for the non-White group (2013-2014). Between-group comparisons of the cohorts showed that non-White participants were more likely to report burnout (OR = 1.36; 95% CI, 1.05–1.57). Greater odds of burnout were found in two studies

Table 3

Racial-Ethnic Identity Predictors of Burnout Outcomes from Included Studies

Authors and Year	Burnout Outcome Measure	Effect Measure	Racial-Ethnic Predictor(s)	Effect Size/Statistics	Interpretation of Results
Agha et al. (2024)	MBI-HSS	Multivariable Logistic Regression (OR)	Asian White Other	1 [Reference] (AOR = 0.77; 95% CI, 0.27–2.20; *p* = 0.629) (AOR = 3.60; 95% CI, 0.36–38.09; *p* = 0.273)	No significant associations were found for White and Other when compared to the Asian group.
Akinleye et al. (2024)	Mini Z	Poisson Regression (RR)	Non-Hispanic White Hispanic Black Other Asian	1 [Reference] (ARR = 0.88; 95% CI, 0.66–1.20) (ARR = 0.78; 95% CI, 0.58–1.10) (ARR = 0.79; 95% CI, 0.58–1.10) (ARR = 0.61; 95% CI, 0.47–0.79; *p*<.001)	Lower burnout risk was associated with the Asian group compared with non-Hispanic White. No other associations were significant.
Bourne et al. (2019)	MBI-HSS-MP	Multivariable Logistic Regression (OR)	White Asian Black Multiracial Other	1 [Reference] (AOR = 0.74; 95% CI, 0.60–0.91) (AOR = 0.73; 95% CI, 0.51–1.02) (AOR = 0.82; 95% CI, 0.58–1.15) (AOR = 2.19; 95% CI, 1.37–3.52)	Lower odds of burnout were found for the Asian group, whereas Other were associated with higher odds. No other associations were significant relative to the White group.
Carthon et al. (2024)	MBI-HSS (EE subscale)	Multi-Level Logistic Regression (OR)	Non-Hispanic White Hispanic	1 [Reference] (AOR = 1.04; 95% CI, 0.83–1.21)	No significant associations were found between the non-Hispanic White and Hispanic groups.
Coombs et al. (2020)	MBI-HSS	Chi-Square	White Asian Indian/South Asian Black/African Latinx Hispanic Other	While there was a higher prevalence of burnout among White participants, no statistically significant relationships were found (*p* = 0.479).	No statistically significant relationships were found between burnout and racial-ethnic identity.
Douglas et al. (2021)	MBI-HSS (EE, DP)	Ordinal Logistic Regression (OR)	Non-Underrepresented: Asian	1 [Reference]	Lower odds of emotional exhaustion and depersonalisation were found for

			White Other		URGs (Black, Hispanic/Latinx, American Indian).
			Underrepresented: Black Hispanic/Latinx American Indian	EE: (AOR = 0.85; 95% CI, 0.71–1.02; *p* = 0.08) DP: (AOR = 0.74; 95% CI, 0.62–0.90; *p* = .0002)	
Dyrbye et al. (2022)	MBI-HSS (EE, DP)	Multivariable Logistic Regression (OR)	Non-Hispanic White Black Hispanic/Latinx Asian/Native Hawaiian /Pacific Islander Indigenous/Other ≥2 Races	1 [Reference] (AOR = 0.81; 95% CI, 0.58–1.14) (AOR = 074; 95% CI, 0.58–0.94) (AOR = 0.67; 95% CI, 0.54–0.79; *p*<.001) (AOR = 0.73; 95% CI, 0.52–1.03) (AOR = 1.27; 95% CI, 0.82–1.96)	Lower odds of burnout were found for non-Hispanic Asian, Native Hawaiian, and Pacific Islanders compared to the non-Hispanic White group in the multivariable analysis.
Garcia et al. (2020)	MBI-HSS (EE, DP)	Multivariable Logistic Regression (OR)	Non-Hispanic White Hispanic/Latinx Black, non-Hispanic Asian, non-Hispanic	1 [Reference] (AOR = 0.63; 95% CI, 0.47–0.86; *p* = 0.004) (AOR = 0.49; 95% CI, 0.30–0.79; *p* = 0.004) (AOR = 0.77; 95% CI, 0.61–0.96; *p* = 0.02)	Lower odds of burnout were found for Asian, Black, and Hispanic/Latinx groups relative to non-Hispanic White. Identifying as Black showed the strongest association.
Khan et al. (2021)	MBI-HSS-MP	Multivariable Logistic Regression (OR)	White Underrepresented: Asian/Pacific Islander South Asian Other	1 [Reference] (AOR = 1.81; 95% CI, 1.28−2.55; *p*<.001)	Low personal accomplishment was more likely to be reported by URGs. No associations between racial-ethnic group and emotional exhaustion or depersonalisation were found.
Lee et al. (2021)	MBI-HSS-MP	Multivariable Linear Regression (*B*)	White Non-White (Total) Non-White (Low PA) Non-White (EE) Non-White (DP)	1 [Reference] (*B* = –0.23; 95% CI, −0.45 to −0.01; *p* = 0.045) (*B* = –0.24; 95% CI, –0.02 to 0.33; *p* = 0.07) (*B* = –0.24; 95% CI, –0.59 to 0.10; *p* = 0.16) (*B* = 0.16; 95% CI –0.62 to –0.01; *p* = 0.045)	Lower overall burnout and depersonalisation on the MBI subscale was found for the non-White group relative to White counterparts.
Lu et al. (2023)	Single-Item (EE, DP)	Multivariable Logistic Regression (OR)	White Underrepresented	1 [Reference] (AOR = 0.81; 95% CI, 0.71−0.93)	URGs (28.9%) showed a lower prevalence of burnout than the White

					group (32.4%). Lower odds of burnout were found for URGs.
Martinez et al. (2022)	MBI-HSS	Multivariable Logistic Regression (OR)	2013-2014 Cohort: White Non-White (Total) Non-White (EE) Non-White (DP) 2019-2020 Cohort: White Non-White (Total)	 1 [Reference] (AOR = 0.60; 95% CI, 0.41–0.86) (AOR = 0.54; 95% CI, 0.35–0.81) (AOR = 0.62; 95% CI, 0.40–0.97) 1 [Reference] (AOR = 0.60; 95% CI, 0.39–0.91)	Lower odds of overall burnout were reported for the non-White groups in both cohorts. In the 2013-2014 cohort, lower rates of emotional exhaustion and depersonalisation were reported on the MBI subscales for the non-White group compared to White.
Merces et al. (2020)	MBI-HSS	Poisson Regression (PR)	Non-Black Black	1 [Reference] (aPR = 0.62; 95% CI, 0.47–0.83)	Higher probability of burnout was found among non-Black participants.
Murray and Chiotu (2024)	ProQol	Multiple Linear Regression (*B*)	White Asian Black Indigenous	1 [Reference] ($B = –2.30$; 95% CI, −4.47 to 3.64; $p = 0.091$)	No significant associations were found for the Asian, Black, and Indigenous groups compared to White.
Nituica et al. (2021)	aMBI	Multiple Linear Regression (*B*)	White Asian Hispanic Black Multiracial/Other	1 [Reference] ($B = −0.27$; $p = 0.001$) ($B = −0.35$; $p = 0.012$) ($B = −0.55$; $p = 0.002$)	Black and Hispanic participants had lower depersonalisation scores and Asian participants reported lower emotional exhaustion. No associations were found for personal accomplishment.
Parker et al. (2025)	MBI-HSS-MP	Multivariable Logistic Regression (OR)	White Asian Black Multiracial Other	1 [Reference] (AOR = 0.97; 95% CI, 0.64–1.48) (AOR = 1.29; 95% CI, 0.67–2.61) (AOR = 1.51; 95% CI, 0.56–5.00) (AOR = 2.48; 95% CI, 1.16–5.91)	Higher odds of burnout were found for the Other racial-ethnic group relative to their White counterparts.
Pillado et al. (2022)	aMBI	Chi-Square	White Asian Hispanic/Latinx Black Other	Around 45.7% of participants reported burnout, although no significant differences were found.	No statistically significant differences were found between the groups.

Robinson et al. (2024)	MBI-HSS-MP	Multi-Level ANOVA (*F*)	White Black Asian/Pacific Islander Hispanic/Other	Significant effect of race on the depersonalisation subscale ($F(3, 146) = 4.2$, $p < .01$) with highest mean scores reported for Asian/Pacific Islanders.	Highest mean depersonalisation scores were found for Asian/Pacific Islanders. Post hoc comparison using Tukey showed Asian/Pacific Islanders met the criteria for high emotional exhaustion along with the lowest personal accomplishment scores.
Sudol et al. (2021)	MBI-HSS	Logistic Regression (OR)	White Underrepresented: Black Hispanic Pacific Islander	1 [Reference] (AOR = 2.08; 95% CI, 1.31−3.30; $p = .002$)	Low personal accomplishment was more likely to be reported among URGs (Black, Hispanic, Pacific Islander) than the White group.
Taylor et al. (2024)	CBI (Work-Related Subscale)	Hierarchical Multiple Regression (*B*)	White Black Asian Hispanic/Latinx Multiracial Other	Significantly less burnout was reported in the Hispanic/Latinx group ($B = -1.18$; $sr^2 = 0.015$; $p < .001$) when compared to the Other racial-ethnic group.	Significantly less burnout was reported among the Hispanic/Latinx group. No associations were found for Asian, Black, and White groups.

Note. Reference Group = 1 [Reference]; MBI-HSS = Maslach Burnout Inventory-Human Services Survey; MBI-HSS (MP) = Maslach Burnout Inventory-Human Services Survey for Medical Personnel; aMBI = Abbreviated Maslach Burnout Inventory; CBI = Copenhagen Burnout Inventory; ProQol = Professional Quality of Life Scale; EE = Emotional Exhaustion; DP = Depersonalisation; PA; Personal Accomplishment; AOR = Adjusted Odds Ratio; ARR = Adjusted Rate Ratio; aPR = Adjusted Prevalence Ratio; CI = Confidence Interval; ANOVA = Analysis of Variance; URG = Underrepresented Group.

among participants identifying as ‘Other’ (Bourne et al., 2019; Parker et al., 2025). For domain-level burnout, URGs (Asian/Pacific Islander, Black, Hispanic, South Asian, Other) were more likely to report low personal accomplishment (Khan et al., 2021; Sudol et al., 2021). Two studies did not find associations between racial-ethnic identity and burnout (Carthon et al., 2024; Murray & Chiotu, 2024). Overall, participants from URGs appear to be at lower risk of burnout relative to White participants, with Asian, Black, and Hispanic/Latinx groups showing consistently reduced burnout rates.

Studies Comparing Burnout Outcomes Across Racial-Ethnic Identities

Seven studies compared burnout outcomes across a range of racial-ethnic identities. These studies comprised samples of doctors or nurses, apart from one which used a combined sample (Robinson et al., 2024). Three studies identified protective effects among American Indian, Black, and Hispanic/Latinx participants, either for overall burnout or depersonalisation when compared with non-Black (Merces et al., 2020) or Asian, White, multiracial, or ‘Other’ groups (Douglas et al., 2021; Taylor et al., 2024). One study revealed higher depersonalisation in Asian/Pacific Islanders participants relative to White, Black, Hispanic, or ‘Other’ groups (Robinson et al., 2024). Three studies did not find associations between racial-ethnic identity and burnout (Agha et al., 2024; Coombs et al., 2020; Pillado et al. 2022). Overall, Asian, Black, and Hispanic/Latinx groups generally reported lower burnout; however, effects varied across subscales and specific racial-ethnic identities.

Racial-Ethnic Identity, Professional Role, and Burnout

Studies Examining Registered Medical Doctors

Fourteen studies examined doctors, of which 12 reported significant findings. Lower odds of overall burnout were most frequently reported for Asian, Black, and Hispanic/Latinx doctors relative to White (e.g., Bourne et al., 2019; Douglas et al., 2021; Dyrbye et al., 2022; Garcia et al., 2020; Nituica et al., 2021). Speciality and training grade (i.e., consultants, speciality doctors, residents) varied across studies, with obstetrics being the most frequently reported speciality. Non-White emergency medicine doctors showed a similar trend of reduced overall burnout, as well as lower depersonalisation scores (Lee et al., 2021; Lu et al., 2023; Martinez et al. 2022). Conversely, doctors in obstetrics services who identified from 'Other' groups displayed greater burnout across various seniority levels (Bourne et al., 2019; Parker et al., 2025). Three studies found no significant associations (Agha et al., 2024; Coombs et al., 2020; Pillado et al., 2022). Collectively, underrepresented doctors seem to be at lower risk of burnout unless identifying with a 'Other' racial-ethnic group, indicating a protective trend against burnout in certain subgroups of doctors.

Studies Examining Registered Nurses

Four studies examined nurses, of which two reported significant findings. Lower burnout was found for Black primary care nurses compared to non-Black nurses (Merces et al., 2020). Conversely, Taylor et al. (2024) identified significantly less burnout among Hispanic/Latinx acute care nurses, whereas Asian, White, Black and multiracial nurses showed no associations with burnout. Remaining studies found no differences among Hispanic or Asian nurses working in acute care and surgical services (Carthon et al., 2024; Murray & Chiotu, 2024) when compared to a White reference group. While substantially less nurse samples were examined, there are evident inconsistencies among Asian groups who showed no associations in nurse studies, whereas these groups often reported lower burnout in doctor samples. This suggests that burnout differences in HCPs may be shaped by factors other than racial-ethnic characteristics alone.

Studies Examining Registered Doctors and Nurses

Two studies examined pooled samples of doctors and nurses, of which both reported significant findings. Across racial-ethnic groups (i.e., Asian/Pacific Islander, Black, Hispanic, White, Other), Asian/Pacific Islanders working in maternity units exhibited worse outcomes with significantly higher depersonalisation scores (Robinson et al., 2024). Conversely, lower overall burnout was associated with Asian primary care doctors and nurses relative to their White counterparts (Akinleye et al., 2024). Subgroup analyses showed that resident doctors were associated with significantly higher emotional exhaustion ($F(3, 146) = 4.0, p < .01$) and depersonalisation ($F(3, 146) = 19.10, p < .001$) than nurses and more experienced doctors (Robinson et al., 2024), whereas doctors and nurses showed no differences in burnout when models were adjusted for demographic and clinical factors in Akinleye et al. (2024). Across studies, the contrasting results among Asian groups based on role are notable, as well as the greater rates of burnout reported in less experienced medical roles.

Racial-Ethnic Identity, Mistreatment, and Burnout

Five studies explored associations between racial-ethnic identity, mistreatment and/or burnout in doctors. In studies that used White or non-Hispanic White as the reference group, doctors from a URG (e.g., Asian, Black, Hispanic/Latinx, Indigenous/Other) were more likely to experience racial discrimination such as derogatory remarks or misidentification of role or ethnicity (Dyrbye et al., 2022; Pillado et al., 2022). Additionally, Dyrbye et al. (2022) found experiences of mistreatment or discrimination were associated with higher odds of burnout among URGs. Conversely, Lu et al. (2023) reported even lower odds of burnout for URGs (e.g., Asian, Black, Multiracial) after controlling for mistreatment. Across racial-ethnic group comparisons, Sudol et al. (2021) found that underrepresented female doctors (i.e., Black, Hispanic, Pacific Islander) who experienced microaggressions (e.g., implicit racial bias) were more likely to report burnout, whereas Agha et al. (2024) found no influence of racial-ethnic identity on discrimination. Results suggest that overt group differences in burnout may not always be observed. However, exposure to mistreatment or discrimination appears to increase burnout risk among URGs in medical doctor populations.

Table 4

Racial-Ethnic Mistreatment Predictors of Burnout Outcomes from Included Studies

Authors and Year	Outcome Measure	Effect Measure	Mistreatment Predictor(s)	Racial-Ethnic Predictor(s)	Effect Size (Adjusted)	Interpretation of Results
Agha et al. (2024)	MBI-HSS Self-Report Questionnaire (Discrimination, Harassment, Stress-Related Factors)	Multivariable Logistic Regression (OR)	Mistreatment	White Asian Other	Prevalence of workplace bullying/harassment (35.6%) and discrimination (30.8%) in the total sample. Multivariate analysis found these factors did not have an impact on burnout: bullying/harassment (OR = 3.75; 95% CI, 0.93−15.12; p = 0.063) and discrimination (OR = 0.70; 95% CI, 0.17−2.94; p = 0.629). Univariate analysis found mistreatment was not influenced by racial-ethnic identity: harassment (R^2 = 0.10; p = 0.31) and discrimination (R^2 = −0.07; p = 0.47).	Exposure to mistreatment did not influence burnout rates and was not impacted by the racial-ethnic identity of doctors.
Dyrbye et al. (2022)	MBI-HSS (EE; DP) Six-Item Mistreatment and Discrimination Questionnaire	Multivariable Logistic Regression (OR)	Mistreatment Discrimination	Non-Hispanic White Underrepresented: Black Hispanic/Latinx Asian/Native Hawaiian/Pacific Islander Indigenous/Other ≥2 Races	All URGs were more likely to experience mistreatment/discrimination compared to non-Hispanic White (e.g., Black: OR = 1.59; 95% CI, 1.13−2.23). These experiences were associated with higher odds of burnout. Odds of burnout were lower for Asian/Native Hawaiian/Pacific Islander groups than non-Hispanic White after controlling for mistreatment/discrimination and other personal and professional factors (OR = 0.67; 95% CI, 0.54−0.79; p<.001).	URGs reported more exposure to workplace mistreatment and discrimination in the past year which was associated with higher burnout rates compared to non-Hispanic White doctors.
Lu et al. (2023)	Single- Item (EE, DP)	Multivariable Logistic Regression (OR)	Mistreatment	White Underrepresented	1 [Reference] (OR = 0.75; 95% CI, 0.65−0.86)	In adjusted models including mistreatment measures, URGs (Black, Asian/Pacific Islander,

	Self-Report Questionnaire (Mistreatment)					Hispanic/Latinx, Native American/Alaska Native, Other/Multiracial) continued to show lower odds of burnout.
Pillado et al. (2022)	aMBI Six-Item Racial Discrimination Questionnaire	Multivariable Logistic Regression (OR)	Racial-Ethnic Discrimination	White Asian Hispanic/Latinx Black Other	1 [Reference] (OR = 6.9; 95% CI, 3.53−13.3; *p*<.001) (OR = 2.7; 95% CI, 0.79−5.36; *p* = 0.06) (OR = 13.6; 95% CI, 4.25−43.40; *p*<.001) (OR = 2.1; 95% CI, 0.79−5.36; *p* = 0.14) Prevalence of burnout was Black (56.3%), Asian (36.3%), Hispanic/Latinx (18.2%), Other (17.1%), and White (9.9%).	Black and Asian doctors reported more racial discrimination compared to other groups. Multivariable analysis showed Black and Asian race to be a risk factor for racial discrimination. The interaction of burnout and racial discrimination was not examined.
Sudol et al. (2021)	MBI-HSS RMAS	Logistic Regression (OR)	Mistreatment	White Underrepresented: Black Hispanic Pacific Islander	Underrepresented females (OR = 2.32; 95% CI, 1.41−3.82; *p*<.001) and males (OR = 1.82; 95% CI, 1.12−2.94; *p* = 0.01) had higher odds of experiencing racial-ethnic microaggression than White male doctors. Underrepresented female doctors (OR = 1.86; 95% CI, 1.03−3.35; *p* = 0.04) who experienced racial-ethnic microaggressions were more likely to report burnout than White male doctors.	Identifying as female and from a URG was identified as a risk factor for racial-ethnic microaggression and these experiences increased burnout risk.

Note. Reference Group = 1 [Reference]; MBI-HSS = Maslach Burnout Inventory-Human Services Survey; aMBI = Abbreviated Maslach Burnout Inventory; EE = Emotional Exhaustion; DP = Depersonalisation; RMAS = Racial Microaggressions Scale; OR = Odds Ratio; CI = Confidence Interval; URGs = Underrepresented Groups.

Outcome Measures

Studies Comparing Racial-Ethnic Identity with Burnout Measures

The MBI-HSS or adapted versions were the most commonly reported measure in 14 studies. Remaining studies used the CBI, ProQol, Mini Z instrument, or single-item MBI tools measuring emotional exhaustion and depersonalisation. The majority of studies using the MBI reported significant associations between racial-ethnic identity and burnout; however, the direction of effect varied across burnout domains. Studies that applied predetermined cut-off scores for the MBI consistently reported lower overall burnout or depersonalisation among Asian, Black, and Hispanic/Latinx groups (Bourne et al., 2019; Dyrbye et al., 2022; Garcia et al., 2020; Merces et al., 2020). Similarly, the Mini Z or single-item tools showed markedly reduced burnout for Asian groups and lower depersonalisation for URGs (Akinleye et al. 2024; Douglas et al. 2021). The aMBI revealed less consistent racial-ethnic differences as White participants displayed higher depersonalisation scores compared with Hispanic and Black groups (Nituica et al., 2021; Pillado et al., 2022), while other domains showed no significant associations.

The ProQol and CBI burnout subscales measured burnout as a continuous outcome. One study used the CBI (Taylor et al., 2024), reporting lower burnout for the Hispanic/Latinx group. No significant associations were found on the ProQol (Murray & Chiotu, 2024). Overall, the relationship between racial-ethnic identity and burnout varied according to the measurement tool and statistical method used to analyse results. Several studies found stronger associations between racial-ethnic identity and the depersonalisation or personal accomplishment MBI subscales. Moreover, multidimensional tools (e.g., MBI-HSS) were more likely to detect significant differences, whereas studies using abbreviated measures and/or dichotomous outcomes typically found no associations. While the majority of MBI studies may partly explain these differences, it is possible binary burnout classifications reduced measurement sensitivity in some studies (e.g., Agha et al., 2024; Carthon et al., 2024; Pillado et al., 2022).

Studies Comparing Racial-Ethnic Identity with Mistreatment Measures

Five studies utilised mistreatment measures. Across studies, the validated and more reliable measures were more likely to detect significant associations between racial-ethnic identity, experiences of mistreatment, and burnout. For example, the RMAS identified greater exposure to mistreatment and odds of burnout among URGs (Sudol et al., 2021), as did the six-item survey used in Dyrbye et al. (2022), which was modified from the Association of American Medical Colleges Graduation Questionnaire national survey. Broader self-report questionnaires, either researcher-developed or based on surveys from similar medical studies, tended to show no associations or inconsistent findings (Agha et al., 2024; Lu et al., 2023; Pillado et al., 2022). Overall, the validated measures should be prioritised over researcher-developed tools, providing increased accuracy and credibility of findings.

Discussion

Summary of Main Findings

This systematic review summarises the variation in burnout across a range of racial-ethnic identities in existing literature, providing a contribution to the current understanding of burnout. A total of 20 observational studies were identified from multiple research databases, examining race and ethnicity as a predictor of burnout among registered doctors and nurses. Five studies measured experiences of mistreatment as a variable alongside burnout, allowing for additional analyses to examine independent and combined effects. Racial-ethnic identity was associated with burnout in the majority of studies; however, there were inconsistencies in the direction of the effect and how burnout was measured. Around 50% of studies identified lower overall and domain-specific (i.e., emotional exhaustion, depersonalisation) burnout risk among URGs (i.e., Asian, Black, Hispanic/Latinx) relative to White groups, as well as across all racial-ethnic groups. However, remaining studies found no differences or reported worse effects for URGs (e.g., Asian, Black, Hispanic/Latinx, Pacific Islander, Other), particularly for reduced personal accomplishment.

Findings suggest that apparent protective effects of identifying from a URG ought to be interpreted cautiously. Role-based inconsistencies in burnout levels and subgroup and subscale variability across studies should be noted. While studies using White doctors as the reference group often reported lower burnout among Asian doctors, analyses comparing all racial-ethnic groups showed greater burnout (e.g., depersonalisation) or no associations for Asian doctors and nurses. These mixed findings likely reflect methodological and contextual factors rather than true differences in burnout risk (Cantor & Mouzon, 2020). Cultural factors may influence differential reporting among URGs, such as norms discouraging emotional expression and disclosure, as well as stronger personal coping strategies appearing to buffer against workplace challenges (e.g., strong work ethic, support networks). For example, URGs showed lower emotional depletion and cynicism in several studies, whereas reduced personal accomplishment in other studies suggest that these groups could still encounter barriers to feeling effective in their roles (Odom et al., 2007).

Burnout measurement is another notable consideration for the mixed findings. Despite well-established measures, these have been developed in Western contexts which may limit the capability of tools to detect differential expressions of burnout among culturally diverse groups (e.g., physical vs. emotional symptoms). In addition, studies comprising HCPs with different levels of seniority may not have accurately captured exposure to workplace stressors. For example, over-representation in high-workload and low-autonomy roles (e.g., less experienced doctors) could have contributed to greater burnout. Smaller samples of URGs when compared with larger White reference groups could have also generated effects strongly shaped by White variance, increasing the risk of false null findings and misleading conclusions among URGs. Finally, preliminary evidence suggests a link between microaggressions and greater burnout among underrepresented doctors, although strong conclusions cannot be made due to limited studies. In sum, the current evidence is more consistent with differences in the expression and shaping of burnout across diverse groups, rather than racial-ethnic identity alone.

Quality of Included Studies

Included studies were all observational, meaning causal relationships could not be established. The overall quality of studies was low to moderate with many studies identified as high risk of sampling bias; however, there was no clear trend regarding the significance of findings based on a low, unclear, or high rating on the MMAT sampling criterion. Reduced statistical power may have restricted the ability to detect true associations in some studies, specifically those with smaller sample sizes of URGs. There was considerable variability in the classification of racial-ethnic identities as several studies did not provide disaggregated data, nor did they compare sample racial-ethnic data with the base rates of these groups in the target populations. Studies seldom provided explanations for this, and as such there is increased risk of misinterpretation (i.e., overestimation or underestimation) of results as observed burnout rates among URGs might not reflect the true prevalence of underrepresented HCPs in local and national healthcare workforces. The interpretability and generalisability of findings across studies are therefore more limited.

The self-report nature of validated burnout tools across studies may have introduced response bias. Studies that applied the full-scale MBI-HSS permitted a more comprehensive assessment of the multidimensional relationship between race, ethnicity, and burnout. Despite this, several studies used aggregated measures that combined multiple forms of mistreatment (e.g., physical harm, racism, sexism), making it difficult to determine the specific drivers of associations. While multivariable models minimised confounding bias by controlling for other influences (e.g., gender, speciality, training grade), this type of modelling can conceal critical associations (Cantor & Mouzon, 2020). For example, the literature indicates greater burnout among underrepresented medical trainees, particularly in renal and emergency medicine (GMC, 2025; Graham-Brown et al., 2021). Moreover, adjusting for speciality or training grade in models could mask the extent of burnout in these groups as occupational factors may moderate how strongly racial-ethnic identity is associated with burnout. Therefore, interpretations should consider the identified methodological issues.

Positioning Within the Current Evidence Base

In line with previous research, this review highlights the complexities of examining burnout across racial-ethnic identities. Comparably, a systematic review of North American studies reported inconsistent burnout levels among an underrepresented in medicine (URiM) population (e.g., physicians, residents, medical students) relative to non-URiM (Lawrence et al., 2021). Two thirds of included studies in the present review were conducted in the US, enabling meaningful comparisons, particularly for Hispanic/Latinx qualified doctors who showed lower burnout in both reviews. National US workforce surveys indicate that Asian, Black, and Hispanic staff make up a third of nurses (Hynson et al., 2022; Smiley et al., 2025) with Asian doctors considered overrepresented compared to general UK and US populations (Association of American Medical Colleges, 2019; Stockton & Warner, 2024). This might partly explain the greater proportion of these specific URGs across study samples along with stronger group-level burnout associations.

The burnout literature acknowledges the demanding roles of doctors and nurses as increasing the risk of adverse outcomes, with experiences of mistreatment posing additional stressors (Bafna et al., 2025). While URGs were observed to report lower burnout in some studies, an emerging trend suggested that underrepresented doctors were disproportionately impacted by mistreatment or discrimination, and these exposures were associated with elevated burnout. This conflicting picture indicates that social and environmental factors may play an important role in burnout risk, particularly the lived experiences of URGs, such as (but not limited to) discrimination and lack of institutional support. Traditional burnout tools may fail to recognise how prolonged exposure to such experiences may increase stress and impact burnout. This aligns with the broader research field showing poorer wellbeing and career outcomes for underrepresented HCPs exposed to greater racial discrimination (Hu et al., 2019), highlighting the critical need to better understand how URGs conceptualise burnout (e.g., exposure to discrimination, perceived inequity, isolation).

Moreover, the conceptualisation of burnout among URGs is limited to the scope of existing measures. While well-established tools (e.g., MBI-HSS, OLBI, CBI) have undergone cross-cultural validation in various countries and languages, they neglect unique experiences (e.g., individual and systemic racism) that have the potential to affect burnout. Such issues could lead to underreporting and misinterpretation of prevalence rates, which is a pre-existing concern for URGs according to trade unions for doctors (BMA, 2022). This provides another plausible explanation for the variability in burnout among URGs, specifically lower burnout rates which contrast with wider literature. However, this review did not directly examine other key factors that may interfere with burnout (e.g., age, gender, setting, speciality). Interpretation of findings should therefore consider the meaningful intersection of racial-ethnic identity and other contextual factors along with how existing burnout tools or within-group variation, such as heterogeneity among specific racial-ethnic minority groups, can shape burnout experiences for individual HCPs.

Strengths and Limitations

A main strength of the review is that the protocol was pre-registered, providing a clear explanation of the planned methodology to minimise reporting bias and prevent unintended duplication (Pieper & Rombey, 2022). Searches were conducted across six major and subject-specific databases (e.g., CINAHL for nursing), increasing the recall precision of relevant literature. While a broad and inclusive search strategy resulted in studies from several countries, it is recognised that the majority were conducted in the US where Hispanic/Latinx groups are identified as the largest minority ethnic group. This could have masked nuanced differences among smaller racial-ethnic groups in some studies and limits generalisability somewhat to other countries. Another strength is that two independent reviewers took part in the screening process and risk of bias assessments. This helped to minimise subjectivity and enabled inter-rater reliability to be established, increasing the overall accuracy and validity of findings (McHugh, 2012). However, the inclusion of studies could have been more consistent through the second reviewer screening all of the titles, abstracts, and full texts, rather than a percentage of the titles and abstracts.

Several notable limitations were identified, including considerable variability in the classification of racial-ethnic identities, outcome measures, and effect size measurements, meaning a meta-analysis was not feasible. The absence of pooled estimates therefore limits the validity and confidence in findings (Campbell et al., 2020). In addition, several studies used race and ethnicity interchangeably and compared groups in this way (e.g., White and Hispanic). While this is considered a pragmatic choice for statistical research, it makes it difficult to discern the role of cultural influences (Flanagin et al., 2021). Inclusive language guidelines were followed in the review (APA, 2019b; British Psychological Society, 2025); however, it is recognised that racial-ethnic terms will continue to evolve, and other preferred definitions may be more appropriate to reduce bias even further. The majority of studies used White HCPs as the reference group, shaping burnout differences among URGs as variations from White experiences, rather than providing insight into between-group variability among racial-ethnic minority groups. The review may therefore underrepresent the subtle differences among URGs, further limiting generalisability to more diverse workforces.

Although it is acknowledged that broad search terms helped to capture a range of studies, none of the studies examined religion as a predictor of outcome which limited exploration of intersectionality. Moreover, the search strategy could have been improved through the use of pairs of concepts (e.g., religion and burnout) and a wider range of synonyms. Additionally, removing predictor as a search term may have identified studies where predictor variables were examined in secondary analyses. The review included peer-reviewed, English language articles and excluded grey literature, which may have introduced response bias as studies with non-significant results could have been missed (Seehra et al., 2023). Finally, the majority of studies compared HCPs from URGs with White reference groups. Findings of the review may therefore be more representative of burnout among White and underrepresented populations in medical and nursing settings, rather than subgroups of racial-ethnic minorities.

Implications for Research and Clinical Practice

The main aim of this review was to identify whether racial-ethnic related factors are associated with burnout among doctors and nurses. Findings suggest differential burnout experiences across racial-ethnic identities. The potential overestimation of a protective effect among URGs (e.g., Asian, Black, Hispanic/Latinx) needs to be considered. It is important to not generalise and make assumptions; however, the differences in lived experiences of URGs may help to explain why certain racial-ethnic minority groups reported lower burnout. Stronger personal resilience and coping strategies (e.g., cultural values, family/community support networks, work ethic) may buffer against some workplace challenges and contribute to reduced burnout risk (Hubbell et al., 2024; Nguyen, 2020). However, it is crucial to note they do not remove the harmful impacts of discrimination which appear to heighten burnout risk in URGs. In addition, the smaller sample sizes for URGs may not fully represent their experiences. This accentuates the need for workplace environments to consider how they can mitigate burnout risk and encourage greater representation of URGs in research.

At the organisational level, senior leadership should promote a culture of psychological safety by implementing anti-racist models of care (e.g., large-scale diversity training) and inclusive recruitment practices (NHS England, 2023). This would provide a more diverse and supportive workplace that could protect against burnout (Bafna et al., 2025). Multidimensional measures should also be prioritised to help detect subtle domain-specific differences and improve the measurement of burnout. Future research should attempt to collect complete racial-ethnic data; provide clear categorical descriptions, rather than absolute or mixed classifications (e.g., non-White); and align statistical analyses with the study's conceptualisation of race and ethnicity to ensure accuracy (Ross et al., 2020). This would increase the interpretability of demographic factors, particularly when examining the intersection of other identities such as gender, social class, or sexuality with burnout (APA, 2019a). Several studies in this review recognised the possible impact of the COVID-19 pandemic on burnout. Future research should therefore attempt to account for any long-term impacts of the pandemic when examining risk factors for burnout, particularly in high-risk care settings (e.g., occupational exposure).

The timeliness of this review is supported by included studies being published within the past six years. Future research would benefit from mixed-methods designs such as focus groups or interviews to understand the intersections of individual (e.g., racism, migration, religion) and occupational factors (e.g., occupational positioning) in burnout experiences among URGs. It is plausible that there is a gap in research regarding the intersection of religion and burnout as no eligible studies were identified based on religiosity, highlighting this as another key research area. Qualitative research could also help to explore the cultural relevance of existing burnout definitions (e.g., interpretation of core domains), providing the initial steps in tailoring individual-level interventions which continue to demonstrate cultural acceptance issues among URGs (Bansal et al., 2022; Zhang et al., 2025). Finally, longitudinal designs, extending cross-sectional approaches, could monitor burnout rates alongside organisational changes (e.g., anti-racism strategies), providing a more robust understanding of the effects of workplace discrimination and the efficacy of targeted support across a range of racial-ethnic identities, healthcare workforces, and settings.

Conclusion

This review highlights variation in burnout experiences across different racial-ethnic identities, providing research and clinical implications for medical and nursing professionals. While the observed trend of lower burnout among URGs is worth acknowledging, this was only found in half of the studies with others reporting worse or no effects. This contrasts with national surveys where underrepresented HCPs often report greater burnout. In addition, preliminary results indicate that URGs are disproportionately impacted by workplace mistreatment, leading to elevated burnout risk. This offers direction for future research, to consider the nuanced differences in burnout expression across racial-ethnic identities and how sociocultural, organisational, and methodological factors may shape variations, rather than racial-ethnic identity alone. Longitudinal designs would enable the assessment of burnout during critical periods of organisational change, to explore transient or lasting impacts and provide stronger conclusions to guide targeted and equitable support. Finally, it is hoped the insights gained support a shift towards more inclusive approaches to addressing burnout across all racial-ethnic identities, aiming to strengthen workforce retention, recruitment, and the overall wellbeing of HCPs.

References

Agha, A., Basu, A., Anwar, E., & Hanif, W. (2024). Burnout among diabetes specialist registrars across the United Kingdom in the post-pandemic era. *Frontiers in Medicine*, 11. https://doi.org/10.3389/fmed.2024.1367103

Akinleye, D., Wu., M., Efferen, L. S., McCauley, S., Allen, A., Bennett, H., Snitkoff, L. S., Cleary, L. M., Bliss, K., Martiniano, R., Wang, S., McNutt, L. A., & Osinaga, A. (2024). Newly acquired burnout during the Coronavirus Disease 2019 (COVID-19) pandemic. A retrospective cohort study on the experiences of New York state primary care clinicians. *Journal of Community Health, 49*(1), 34-45. https://doi.org/10.1007/s10900-023-01247-z

American Psychological Association. (2019a). *APA guidelines on race and ethnicity in psychology*. https://www.apa.org/about/policy/summary-guidelines-race-ethnicity

American Psychological Association. (2019b). *Racial and ethnic identity.* https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/racial-ethnic-minorities

Bafna, T., Malik, M., Choudhury, M. C., Hu, W., Weston, C. M., Weeks, K. R., Connors, C., Burhanullah, M. H., Everly, G., Michtalik, H. J., & Wu, A. W. (2025). Systematic review of well-being interventions for minority healthcare workers. *Frontiers in Medicine, 12*. Article e1531090. https://doi.org/10.3389/fmed.2025.1531090

Bansal, N., Karlsen, S., Sashidharan, S. P., Cohen, R., Chew-Graham, C. A., & Malpass, A. (2022). Understanding ethnic inequalities in mental healthcare in the UK: A meta-ethnography. *PLoS Medicine, 19*(12), Article e1004139. https://doi.org/10.1371/journal.pmed.1004139

Bourne, T., Shah, H., Falconieri, N., Timmerman, D., Lees, C., Wright, A., Lumsden, M. A., Regan, L., & Van Calster, B. (2019). Burnout, well-being and defensive medical practice among obstetricians and gynaecologists in the UK: cross-sectional survey study. *BMJ Open, 9*(11), Article e030968. https://doi.org/10.1136/bmjopen-2019-030968

Bramer, W. M., Milic, J., & Mast, F. (2017). Reviewing retrieved references for inclusion in systematic reviews using EndNote. *Journal of the Medical Library Association, 105*(1), 84-87. https://doi.org/10.5195/jmla.2017.111

British Medical Association (2022). *Delivering racial equality in medicine.* [20220228]. https://www.bma.org.uk/advice-and-support/equality-and-diversity-guidance/race-equality-in-medicine/delivering-racial-equality-in-medicine

British Medical Association. (2024, June 28). *Doctors' titles explained.* https://www.bma.org.uk/advice-and-support/international-doctors/life-and-work-in-the-uk/toolkit-for-doctors-new-to-the-uk/doctors-titles-explained

British Psychological Society (2025). *Inclusive language guidance.* https://cms.bps.org.uk/sites/default/files/2025-06/Inclusive%20Language%20Guidance.pdf

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., Hartmann-Boyce, J., Ryan, R., Shepperd, S., Thomas, J., Welch, V., & Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ, 368*, Article l6890. https://doi.org/10.1136/bmj.l6890

Cantor, J. C., & Mouzon, D. M. (2020). Are Hispanic, Black, and Asian physicians truly less burned out than white physicians? individual and institutional considerations. *JAMA Network Open, 3*(8), Article e2013099. https://doi.org/10.1001/jamanetworkopen.2020.13099

Carthon, J. M. B., Brom, H., Maye., A., Iroegbu, C., Gil., J., Rizzo, J., Amenyedor, K., Montavlo, W., & Villarruel, A. M. (2024). Burnout and psychological distress among

Hispanic nurses across Illinois and New York hospitals: Implications for structural and person-centred solutions. *Nursing Outlook, 72*(5), Article 102234. https://doi.org/10.1016/j.outlook.2024.102234.

Carthon, J. M. B., Travers, J. L., Hounshell, D., Udoeyo, I., & Chittams, J. (2021). Disparities in nurse job dissatisfaction and intent to leave implications for retaining a diverse workforce. *The Journal of Nursing Administration, 51*(6), 310-317. https://doi.org/10.1097/NNA.0000000000001019

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Coombs, D. M., Lanni, M. A., Fosnot, J., Patel., A., Korentager, R., Lin, I. C., & Djohan, R. (2020). Professional burnout in United States plastic surgery residents: Is it a legitimate concern? *Aesthetic Surgery Journal, 40*(7), 802–810. https://doi.org/10.1093/asj/sjz281

Delgadillo, J., Saxon, D., & Barkham, M. (2018). Associations between therapists' occupational burnout and their patients' depression and anxiety treatment outcomes. *Depression and Anxiety, 35*(9), 844-850. https://doi.org/10.1002/da.22766

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*(3), 499-512. https://doi.org/10.1037/0021-9010.86.3.499

Douglas, M., Coman, E., Eden, A. R., Abiola, S., & Grumbach, K. (2021). Lower likelihood of burnout among family physicians from underrepresented racial-ethnic groups. *Annals of Family Medicine, 19*(4), 342-350. https://doi.org/10.1370/afm.2696

Dyrbye, L. N., West, C. P., Sinsky, C. A., Trockel, M., Tutty, M., Satele, D., Carlasare, L., & Shanafelt, T. (2022). Physicians' experiences with mistreatment and discrimination by patients, families, and visitors and association with burnout. *JAMA Network Open, 5*(5), Article e2213080. https://doi.org/10.1001/jamanetworkopen.2022.13080

Flanagin, A., Frey, T., & Christiansen, S. L. (2021). Updated guidance on the reporting of race and ethnicity in medical and science journals. *JAMA, 326*(7), 621-627. https://doi.org/doi:10.1001/jama.2021.13304

Freudenberger, H. J. (1975). The staff burn-out syndrome in alternative institutions. *Psychotherapy: Theory, Research & Practice, 12*(1), 73–82. https://doi.org/10.1037/h0086411

Garcia, L. C., Shanafelt, T. D., West, C. P., Sinsky, C. A., Trockel, M. T., Nedelec, L., Maldonado, Y. A., Tutty, M., Dyrbye, L. N., & Fassiotto, M. (2020). Burnout, depression, career satisfaction, and work-life integration by physician race/ethnicity. *JAMA Network Open, 3*(8), Article e2012762. https://doi.org/10.1001/jamanetworkopen.2020.12762

García, R. G., Sangregorio, M. A., & Sánchez, M. L. L. (2019). Factorial validity of the Maslach Burnout Inventory-Human Services Survey (MBI-HSS) in a sample of Spanish social workers. *Journal of Social Service Research, 45*(2), 207-219. https://doi.org/10.1080/01488376.2018.1480549

General Medical Council. (2025). *National training survey 2025 results.* [GMC/NTS2025/0725]. https://www.gmc-uk.org/education/how-we-quality-assure-education-and-training/evidence-data-and-intelligence/national-training-surveys

Getie, A., Ayenew, T., Amlak, B. T., Gedfew, M., Edmealem, A., & Kebede, W. M. (2025). Global prevalence and contributing factors of nurse burnout: an umbrella review of systematic review and meta-analysis. *BMC Nursing, 24,* Article 596. https://doi.org/10.1186/s12912-025-03266-8

Graham-Brown, M. P., Beckwith, H. K., O'Hare, S., Trewartha, D., Burns, A., & Carr, S. (2021). Impact of changing medical workforce demographics in renal medicine over 7 years: Analysis of GMC national trainee survey data. *Clinical Medicine, 21*(4), e363-e370. https://doi.org/10.7861/clinmed.2020-1065

Hall, L. H., Johnson, J., Watt, I., Tsipa, A., & O'Connor, D. B. (2016). Healthcare staff wellbeing, burnout, and patient safety: a systematic review. *PLoS ONE, 11*(7), Article e0159015. https://doi.org/10.1371/journal.pone.0159015

Harding, A., Ooi, J., Stinton, M., Russell, R., & Parke, S. (2025). Levels of stress and burnout in trainee and qualified NHS psychological professionals: A systematic review and narrative synthesis. *Mental Health & Prevention, 40*, Article 200453. https://doi.org/10.1016/j.mhp.2025.200453

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2024). *Cochrane Handbook for Systematic Reviews of Interventions* (Version 6.5). Cochrane. https://www.cochrane.org/handbook

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285-291. https://doi.org/10.3233/EFI-180221

Hu, Y., Ellis, R. J., Hewitt, D. B., Yang, A. D., Cheung, E. O., Moskowitz, J. T., Potts, J. R., Buyske, J., Hoyt, D. B., Nasca, T. J., & Bilimoria, K. Y. (2019). Discrimination, abuse, harassment, and burnout in surgical residency training. *New England Journal of Medicine, 381(*18), 1741-1752. https://doi.org/10.1056/NEJMsa1903759

Hubbell, S. L., Young, S. E., Duea, S. R., & Prentice, C. R. (2024). Identifying protective factors related to burnout, moral injury, and resilience of registered nurses: An exploratory analysis. *Mental Health Science, 2*, Article e71. https://doi.org/10.1002/mhs2.71

Hynson, E., Bloomer, J., Samson, Z., Price, K., Tran, D., & Muench, U. (2022). Workforce trends of underrepresented minority nurses in the United States over the last decade: progress towards equal representation? *Policy Politics & Nursing Practice, 23*(4), 215-227. https://doi.org/10.1177/15271544221118319

Karuna, C., Palmer, V., Scott, A., & Gunn, J. (2022). Prevalence of burnout among GPs: a systematic review and meta-analysis. *British Journal of General Practice 72*(718), e316-e324. https://doi.org/10.3399/BJGP.2021.0441

Khan, N., Palepu, A., Dodek, P., Salmon, A., Leitch, H., Ruzycki, S., Townson, A., & Lacaille, D. (2021). Cross-sectional survey on physician burnout during the COVID-19 pandemic in Vancouver, Canada: the role of gender, ethnicity and sexual orientation. *BMJ Open, 11*(5), Article e050380. https://doi.org/10.1136/bmjopen-2021-050380

Kim, J. G., Lett, E., Boscardin, C. K., Hauer, K. E., Chen, I. L., Henderson, M. C., Hogan, S. O., Yamazaki, K., Burk-Rafel, J., Fancher, T., Nguyen, M., Holmboe, E. S., McDade, W., & Boatright, D. H. (2025). Community racial and ethnic representation among physicians in us internal medicine residency programs. JAMA Network Open, 8(1), Article e2457310. https://doi.org/10.1001/jamanetworkopen.2024.57310

Kinman, G., Dovey, A., & Teoh. K. (2023). *Burnout in healthcare: risk factors and solutions.* https://www.som.org.uk/sites/som.org.uk/files/Burnout_in_healthcare_risk_factors_and_solutions_July2023_0.pdf

Kristensen, T. S., Borritz, M., Villadsen, E., & Christensen, K. B. (2005). The Copenhagen Burnout Inventory: A new tool for the assessment of burnout. *Work & Stress, 19*(3), 192-207. https://doi.org/10.1080/02678370500297720

Lawrence, J. A., Davis, B. A., Corbette, T., Hill, E. V., Williams, D. R., & Reede, J. Y. (2021). Racial/ethnic differences in burnout: a systematic review. *Journal of Racial and Ethnic Health Disparities, 9*(1), 257-269. https://doi.org/10.1007/s40615-020-00950-0

Lee, D. H., Reasoner, K., Lee, D., Davidson, C., Pennings, J. S., Blazar, P. E., Frick, S. L., Kelly, A. M., LaPorte, D. M., Lese, A. B., Mercer, D. M., Ring, D., Salazar, D. H., Scharschmidt, T. J., Snoddy, M. C., Strauch, R. J., Tuohy, C. J., & Wongworawat, M. D. (2021). Is grit associated with burnout and well-being in orthopaedic resident and faculty physicians? A multi-institution longitudinal study across training levels. *Clinical Orthopaedics and Related Research, 479*(12), 2576-2586. https://doi.org/10.1097/CORR.0000000000001987

Li, L. Z., Yang, P., Singer, S. J., Pfeffer, J., Mathur, M. B., & Shanafelt, T. (2024). Nurse burnout and patient safety, satisfaction, and quality of care. *JAMA Network Open*, *7*(11), Article e2443059. https://doi.org/10.1001/jamanetworkopen.2024.43059

Lin, C., Alimoradi, Z., Griffiths, M. D., & Pakpour, A. H. (2022). Psychometric properties of the Maslach Burnout Inventory for Medical Personnel (MBI-HSS-MP). Heliyon, 8(2), Article e08868. https://doi.org/10.1016/j.heliyon.2022.e08868

Lu, D. W., Zhan, T., Bilimoria, K. Y., Reisdorff, E. J., Barton, M. A., Nelson, L. S., Beeson, M. S., & Lall, M. D. (2023). Workplace mistreatment, career choice regret, and burnout in emergency medicine residency training in the United States. *Annals of Emergency Medicine, 81*(6), 706-714. https://doi.org/10.1016/j.annemergmed.2022.10.015

Macaron, M. M., Segun-Omosehin, O. A., Matar, R. H., Beran, A., Nakanishi, H., Than, C. A., & Abulseoud, O. A. (2023). A systematic review and meta-analysis on burnout in physicians during the COVID-19 pandemic: A hidden healthcare crisis. *Frontiers in Psychiatry, 13*, e1071397. https://doi.org/10.3389/fpsyt.2022.1071397

Martinez, E. C., Hasbun, P. E. G., Vargas, V. P. S. García-González, O. Y., Madera, M. D. F., Capistrán, D. E. R., Carmona, T. C., Cruz, C. S., & Hooper, C. T. (2025). A comprehensive guide to conduct a systemic review and meta-analysis in medical research. *Medicine, 104*(33), Article e41868. https://doi.org/10.1097/MD.0000000000041868

Martinez, K. A., Sullivan, A. B., Linfield, D. T., Shaker, V., Yu, P. C., & Rothberg, M. B. (2022). Change in physician burnout between 2013 and 2020 in a major health system. *Southern Medical Journal, 115*(8), 645-650. https://doi.org/10.14423/SMJ.0000000000001428

Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior, 2*(2), 99-113. https://doi.org/10.1002/job.4030020205

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica, 22*(3), 276-282. https://doi.org/10.11613/BM.2012.031

Merces, M. C., Coelho, J. M. F., Lua, I., De Souza e Silva, D., Gomes, A. M. T., Erdmann, A. L., De Oliveira, D. C., Lago, S. B., Santana, A. I. C., Da Silva, D. A. R., Servo, M. L. S., Sobrinho, C. L. N., Marques, S. C., Figueiredo, V. P., Peres, E. M., De Souza, M. C., França, L. C. M., Maciel, D. M. C., Peixoto, Á. R. S.,... Júnior, A. D. (2020). Prevalence and factors associated with burnout syndrome among primary health care nursing professionals: A cross-sectional study. *International Journal of Environmental Research and Public Health, 17*(2). https://doi.org/doi:10.3390/ijerph17020474

Moola, S., Munn, Z., Sears, K., Sfetcu, R., Currie, M., Lisy, K., Tufanaru, C., Qureshi, R., Mattis, P., & Mu, P. (2015). Conducting systematic reviews of association (etiology): The Joanna Briggs Institute's approach. *International Journal of Evidence-Based Healthcare, 13*(3), 163-169. https://doi.org/10.1097/XEB.0000000000000064

Mosley, T. J., Zajdel, R. A., Alderete, E., Clayton, J. A., Heidari, S., Pérez-Stable, E. J., Salt, K., & Bernard, M. A. (2025). Intersectionality and diversity, equity, and inclusion in the healthcare and scientific workforces. *The Lancet Regional Health – Americas, 41*, Article 100973. https://doi.org/10.1016/j.lana.2024.100973

Murray, D. D., & Chiotu, L. B. (2024). How mental health nurses report their compassion fatigue and compassion satisfaction: A cross-sectional study and the implications for healthcare leaders. *Issues in Mental Health Nursing, 45*(5), 506-519. https://doi.org/10.1080/01612840.2024.2322045

Nash, S (2025, July 31). *From burnout to belief: reflections on reforming the NHS from within.* The King's Fund. https://www.kingsfund.org.uk/insight-and-analysis/long-reads/burnout-belief-reforming-the-nhs

National Health Service Employers (2025, March 14). *NHS Staff Survey 2024: health and wellbeing overview.* https://www.nhsemployers.org/articles/nhs-staff-survey-2024-health-and-wellbeing-overview

National Health Service England. (2023, February 22). *NHS Workforce Race Equality Standard (WRES) 2022 data analysis report for NHS trusts.* https://www.england.nhs.uk/long-read/nhs-workforce-race-equality-standard-wres2022-data-analysis-report-for-nhs-trusts/

Nguyen, A. W. (2020). Religion and mental health in racial and ethnic minority populations: a review of the literature. *Innovation in Aging, 4*(5). https://doi.org/10.1093/geroni/igaa035

Nituica, C., Bota, O. A., Blebea, J., Cheng, C. I., & Slotman, G. J. (2021). Factors influencing resilience and burnout among resident physicians - a national survey. *BMC Medical Education, 21*(1), Article 514. https://doi.org/10.1186/s12909-021-02950-y

Odei, B., & Chino, F. (2021). Incidence of burnout among female and minority faculty in radiation oncology and medical oncology. *International Journal of Radiation Oncology, Biology, Physics, 111*(3), S28-S29. https://doi.org/10.1016/j.ijrobp.2021.07.092

Odom, K. L., Roberts, L. M., Johnson, R. L., & Cooper, L. A. (2007). Exploring obstacles to and opportunities for professional success among ethnic minority medical students. *Academic Medicine, 82*(2), 146-153. https://doi.org/10.1097/ACM.0b013e31802d8f2c

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLoS Medicine*, *18*(3), Article e1003583. https://doi.org/10.1371/journal.pmed.1003583

Parker, N., Falconieri, N., Shah, H., Lees, C., Timmerman, D., Calster, B., & Bourne, T. (2025). Burnout, wellbeing and defensive medical practice in obstetricians and gynecologists in the UK before and after the COVID pandemic: A repeated cross-sectional survey study. *Acta Obstetricia et Gynecologica Scandinavica, 104*(9), 1683-1693. https://doi.org/10.1111/aogs.70006

Pieper, D., & Rombey, T. (2022). Where to prospectively register a systematic review. *Systematic Reviews, 11*(1), Article 8. https://doi.org/10.1186/s13643-021-01877-1

Pillado, E. B., Li, R. D., Eng, J. S., Chia, M. C., Conway, A., Gomez-Sanchez, C., Shaw, P., Sheahan, M. G., Bilimoria, K. Y., Hu, Y., & Coleman, D. M. (2022). Persistent racial discrimination among vascular surgery trainees threatens wellness. *Journal of Vascular Surgery, 77*(1), 262-268. https://doi.org/10.1016/j.jvs.2022.09.011

Poghosyan, L., Aiken, L. H., & Sloane, D. M. (2009). Factor structure of the Maslach Burnout Inventory: An analysis of data from large scale cross-sectional surveys of nurses from eight countries. International *Journal of Nursing Studies, 46*(7), 894-902. https://doi.org/10.1016/j.ijnurstu.2009.03.004

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme.* Lancaster University.

Reis, D., Xanthopoulou, D., & Tsaousis, I. (2015). Measuring job and academic burnout with the Oldenburg Burnout Inventory (OLBI): Factorial invariance across samples and countries. *Burnout Research, 2*(1), 8-18. https://doi.org/10.1016/j.burn.2014.11.001

Robinson, K. A., Atlas, R. O., Storr, C. L., Gaitens, J. M., Blanchard, M., & Ogbolu, Y. (2024). Burnout among nurses, midwives, and physicians in maternity care exposed to traumatic childbirth events. *MCN: The American Journal of Maternal/Child Nursing, 49*(6), 332-340. https://doi.org/10.1097/nmc.0000000000001046

Rolewicz, L., Palmer, B., & Lobont, C. (2024, February 7). *The NHS workforce in numbers.* https://www.nuffieldtrust.org.uk/resource/the-nhs-workforce-in-numbers

Ross, P. T., Hart-Johnson, T., Santen, S. A., & Zaidi, N. L. B. (2020). Considerations for using race and ethnicity as quantitative variables in medical education research. *Perspectives on Medical Education, 9*(5), 318-323. https://doi.org/10.1007/s40037-020-00602-3

Rotenstein, L. S., Torre, M., Ramos, M. A., Rosales, R. C., Guille, C., Sen, S., & Mata, D. A. (2018). Prevalence of burnout among physicians: a systematic review. *JAMA, 320*(11), 1131-1150. https://doi.org/10.1001/jama.2018.12777

Royal College of Nursing (2025, November 17). *RCN reports 116% rise in members facing racist abuse.* https://www.rcn.org.uk/news-and-events/news/rcn-reports-116-percent-rise-in-members-facing-racist-abuse

Salvagioni, D. A. J., Melanda, F. N., Mesas, A. E., González, A. D., Gabani, F. L., & De Andrade, S. M. (2017). Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies. *PLoS ONE, 12*(10), 1–29. https://doi.org/10.1371/journal.pone.0185781

Seehra, J., Khraishi, H., & Pandis, N. (2023). Studies with statistically significant effect estimates are more frequently published compared to non-significant estimates in oral health journals. *BMC Medical Research Methodology, 23*, Article 6. https://doi.org/10.1186/s12874-022-01795-3

Shoman, Y., Marca, S. C., Bianchi, R., Godderis, L., van der Molen, H. F., & Canu, I. G. (2021). Psychometric properties of burnout measures: a systematic review. *Epidemiology and Psychiatric Sciences*, *30*, Article e8. https://doi.org/10.1017/S2045796020001134

Smiley, R. A., Kaminski-Ozturk, N., Reid, M., Burwell, P., Oliveria, C. M., Shobo, Y., Allgeyer, R. L., Zhong, E., O'Hara, C., Volk, A., & Martin, B. (2025). The 2024 national nursing workforce survey. *Journal of Nursing Regulation, 16*(1). https://doi.org/10.1016/S2155-8256(25)00047-X

Soares, J. P., Lopes, R. H., de Souza Mendonça, P. B., Silva, C. R. D. V., Rodrigues, C. C. F. M., & de Castro, J. L. (2023). Use of the Maslach Burnout Inventory among public health care professionals: scoping review. *JMIR Mental Health, 10*, Article e44195. https://doi.org/10.2196/44195

Stamm, B. H. (2010). *The concise ProQol manual* (2nd ed.). Pocatello. https://proqol.org/proqol-manual

Stockton, I., & Warner, M. (2024). *Ethnic diversity of NHS doctors.* (Report No. R294). Institute for Fiscal Studies. https://ifs.org.uk/publications/ethnic-diversity-nhs-doctors

Sudol, N. T., Guaderrama, N. M., Honsberger, P., Weiss, J., Li, Q., & Whitcomb, E. L. (2021). Prevalence and nature of sexist and racial/ethnic microaggressions against surgeons and anesthesiologists. *JAMA Surgery, 156*(5), Article e210265. https://doi.org/10.1001/jamasurg.2021.0265

Syahir, A. N. A., Abidin, M. S. Z., Sa'ari, C. Z., & Rahman, M. Z. A. (2025). Workplace spirituality and its impact on employee well-being: a systematic literature review of global evidence. *Journal of Religion and Health, 64*(5), 3313-3345. https://doi.org/10.1007/s10943-025-02350-2

Taylor, E. J., Morrell, H. E. R., Ada, H. M., Dupée, C., Jordan, M., Radovich, P., & Boyd, K. C. (2024). Do nurse spiritual responses to trauma explain nurse outcomes? A correlational observational study. *Nursing Outlook, 72*(6), Article 102296. https://doi.org/10.1016/j.outlook.2024.102296

Webster, R. K., Liu, R., Karimullina, K., Hall, I., Amlôt, R., & Rubin, G. J. (2019). A systematic review of infectious illness Presenteeism: prevalence, reasons, and risk factors. *BMC Public Health, 19*(1). https://doi.org/10.1186/s12889-019-7138-x

West, C. P., Dyrbye, L. N., & Shanafelt, T. D. (2018). Physician burnout: contributors, consequences and solutions. *Journal of Internal Medicine, 283*(6), 516-529. https://doi.org/10.1111/joim.12752

Wilbur, K., Snyder, C., Essary, A. C., Reddy, S., Will, K. K., & Saxon, M. (2020). Developing workforce diversity in the health professions: a social justice perspective. *Health Professions Education, 6*(2), 222-229. https://doi.org/10.1016/j.hpe.2020.01.002

Woo, T., Ho, R., Tang, A., & Tam, W. (2020). Global prevalence of burnout symptoms among nurses: A systematic review and meta-analysis. *Journal of Psychiatric Research, 123*, 9-20. https://doi.org/10.1016/j.jpsychires.2019.12.015

World Health Organization. (2019, May 28). *Burn-out an "occupational phenomenon": International Classification of Diseases.* https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases

Zhang, Z., Sun, S., Moradbakhti, L., Hall, A., Mougenot, C., Chen, J., & Calvo, R. A. (2025). Health care professionals' engagement with digital mental health interventions in the United Kingdom and China: Mixed methods study on engagement factors and design implications. *JMIR Mental Health*, *12*, Article e67190. https://doi.org/10.2196/67190

Appendices

Appendix A

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Page 1
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	Page 2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Page 7
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Page 7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Page 8-9
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Page 9-10
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Page 9-10
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Page 10
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Page 10-11
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Page 10-11
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Page 10-11
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Page 11
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Page 23-26; 30-31
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Page 11-12
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Page 11-12; 21
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Page 10-11
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Page 11-12
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	N/A
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	N/A
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Page 11

Section and Topic	Item #	Checklist item	Location where item is reported
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Page 11-12
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Page 12-13
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Page 12-13
Study characteristics	17	Cite each included study and present its characteristics.	Page 14-19
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Page 20-21; 65-66
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Page 21-33
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Page 21-33; 34-35
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	N/A
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	N/A
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	N/A
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Page 20-21; 34-35
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Page 34-35
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Page 35-37
	23b	Discuss any limitations of the evidence included in the review.	Page 34-35
	23c	Discuss any limitations of the review processes used.	Page 37-38
	23d	Discuss implications of the results for practice, policy, and future research.	Page 38-41
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Page 8
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Page 8
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	N/A
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page ii
Competing interests	26	Declare any competing interests of review authors.	Page ii
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Page ii

Appendix B

Search Strategy

CINAHL

TX ((ethnic* OR "ethnicity" OR rac* OR religio* OR faith OR cultur* OR racism OR discriminat* OR "global majority" OR minorit* OR diversity OR inequality OR disparit*) AND ("risk factor" OR associat* OR correlat* OR predict* OR "determinants" OR caus*) AND ("cross sectional study" OR longitudinal OR cohort OR "randomised controlled trial" OR RCT OR quasi OR observational OR "non-experimental") AND (burnout OR "burnout syndrome" OR exhaust* OR disengage* OR depersonal* OR "work* stress" OR "occupational stress" OR "maslach burnout inventory" OR "copenhagen burnout inventory" OR "oldenburg burnout inventory") AND ("healthcare professionals" OR "healthcare worker" OR nurs* OR doctor* OR physician* OR "medical staff" OR clinician* OR "registered nurse" OR "nursing staff"))

Medline (via Ovid)

((ethnic* OR "ethnicity" OR rac* OR religio* OR faith OR cultur* OR racism OR discriminat* OR "global majority" OR minorit* OR diversity OR inequality OR disparit*) AND ("risk factor" OR associat* OR correlat* OR predict* OR "determinants" OR caus*) AND ("cross sectional study" OR longitudinal OR cohort OR "randomised controlled trial" OR RCT OR quasi OR observational OR "non-experimental") AND (burnout OR "burnout syndrome" OR exhaust* OR disengage* OR depersonal* OR "work* stress" OR "occupational stress" OR "maslach burnout inventory" OR "copenhagen burnout inventory" OR "oldenburg burnout inventory") AND ("healthcare professionals" OR "healthcare worker" OR nurs* OR doctor* OR physician* OR "medical staff" OR clinician* OR "registered nurse" OR "nursing staff")).af.

PsycINFO / PsycARTICLES (via Ovid)

((ethnic* OR "ethnicity" OR rac* OR religio* OR faith OR cultur* OR racism OR discriminat* OR "global majority" OR minorit* OR diversity OR inequality OR disparit*) AND ("risk factor" OR associat* OR correlat* OR predict* OR "determinants" OR caus*) AND ("cross sectional study" OR longitudinal OR cohort OR "randomised controlled trial" OR RCT OR quasi OR observational OR "non-experimental") AND (burnout OR "burnout syndrome" OR exhaust* OR disengage* OR depersonal* OR "work* stress" OR "occupational stress" OR "maslach burnout inventory" OR "copenhagen burnout inventory" OR "oldenburg burnout inventory") AND ("healthcare professionals" OR "healthcare worker" OR nurs* OR doctor* OR physician* OR "medical staff" OR clinician* OR "registered nurse" OR "nursing staff").ab,hw,id,ti.

Scopus

TITLE-ABS-KEY ((ethnic* OR "ethnicity" OR rac* OR religio* OR faith OR cultur* OR racism OR discriminat* OR "global majority" OR minorit* OR diversity OR inequality OR disparit*) AND ("risk factor" OR associat* OR correlat* OR predict* OR "determinants" OR caus*) AND ("cross sectional study" OR longitudinal OR cohort OR "randomised controlled trial" OR RCT OR quasi OR observational OR "non-experimental") AND (burnout OR "burnout syndrome" OR exhaust* OR disengage* OR depersonal* OR "work* stress" OR "occupational stress" OR "maslach burnout inventory" OR "copenhagen burnout inventory" OR "oldenburg burnout inventory") AND ("healthcare professionals" OR "healthcare worker" OR nurs* OR doctor* OR physician* OR "medical staff" OR clinician* OR "registered nurse" OR "nursing staff"))

Web of Science

((((TS=(ethnic* OR “ethnicity” OR rac* OR religio* OR faith OR cultur* OR racism OR discriminat* OR “global majority” OR minorit* OR diversity OR inequality OR disparit*)) AND TS=("risk factor" OR associat* OR correlat* OR predict* OR “determinants” OR caus*)) AND TS=("cross sectional study" OR longitudinal OR cohort OR “randomised controlled trial” OR RCT OR quasi OR observational OR “non-experimental”)) AND TS=(burnout OR “burnout syndrome” OR exhaust* OR disengage* OR depersonal* OR “work* stress” OR “occupational stress” OR “maslach burnout inventory” OR “copenhagen burnout inventory” OR “oldenburg burnout inventory”)) AND TS=("healthcare professionals" OR “healthcare worker” OR nurs* OR doctor* OR physician* OR “medical staff” OR clinician* OR “registered nurse” OR “nursing staff”)

Appendix C

MMAT (2018): Quantitative Non-Randomised Checklist

MIXED METHODS APPRAISAL TOOL (MMAT) VERSION 2018

User guide

Prepared by
Quan Nha HONG[a], Pierre PLUYE[a,], Sergi FÀBREGUES[b], Gillian BARTLETT[a], Felicity BOARDMAN[c], Margaret CARGO[d], Pierre DAGENAIS[e], Marie-Pierre GAGNON[f], Frances GRIFFITHS[c], Belinda NICOLAU[a], Alicia O'CATHAIN[g], Marie-Claude ROUSSEAU[h], & Isabelle VEDEL[a]

[a]McGill University, Montréal, Canada; [b]Universitat Oberta de Catalunya, Barcelona, Spain; [c]University of Warwick, Coventry, England; [d]University of Canberra, Canberra, Australia; [e]Université de Sherbrooke, Sherbrooke, Canada; [f]Université Laval, Québec, Canada; [g]University of Sheffield, Sheffield, England; [h]Institut Armand-Frappier Research Centre, Laval, Canada

McGill | Department of Family Medicine | Département de médecine de famille
Academic excellence and innovation in care, teaching and research
Innovation et excellence académique dans les soins, l'enseignement et la recherche

Last update: August 1st, 2018

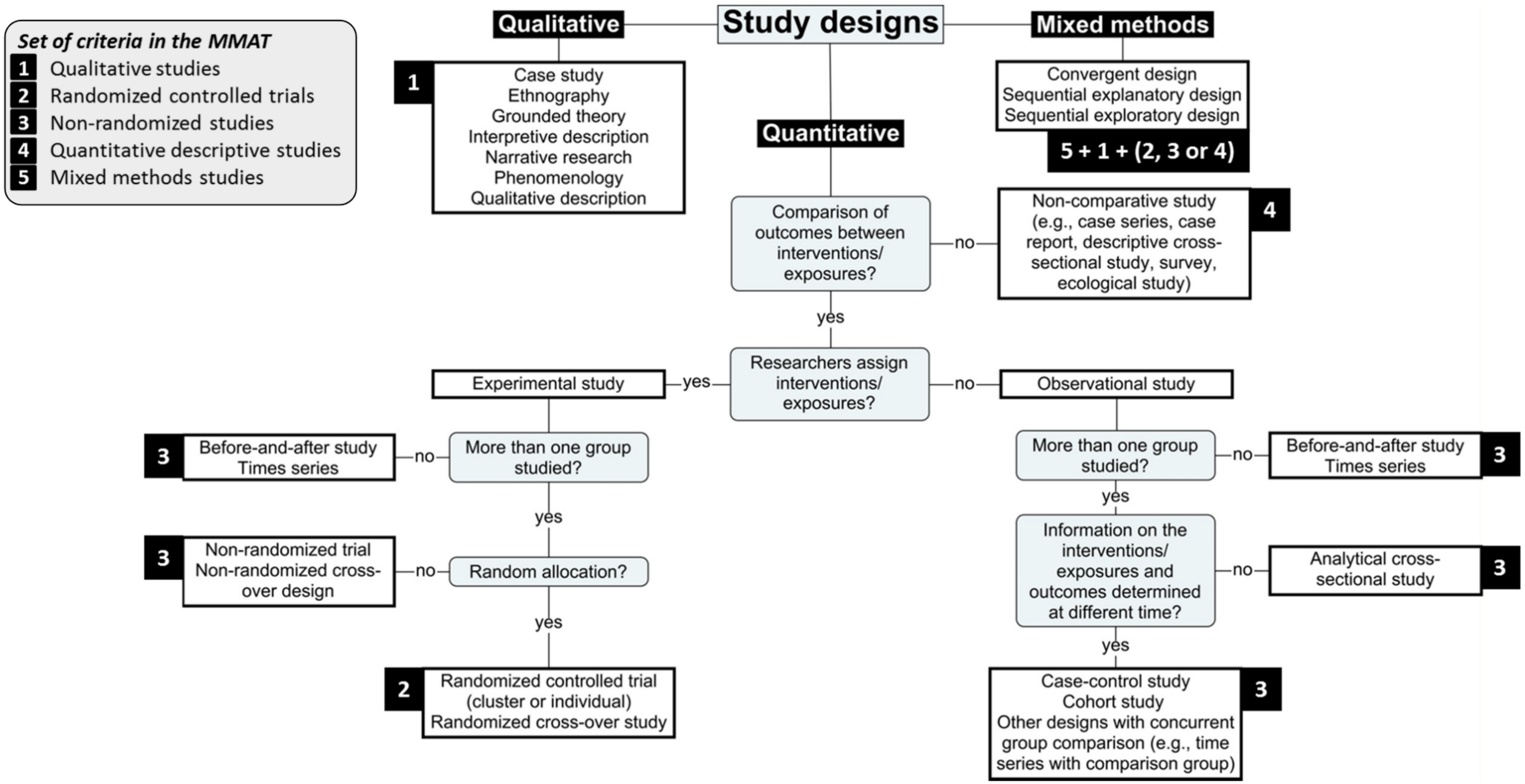
Part I: Mixed Methods Appraisal Tool (MMAT), version 2018

Category of study designs	**Methodological quality criteria**	**Responses**			
		Yes	No	Can't tell	Comments
Screening questions (for all types)	S1. Are there clear research questions?				
	S2. Do the collected data allow to address the research questions?				
	Further appraisal may not be feasible or appropriate when the answer is 'No' or 'Can't tell' to one or both screening questions.				
1. Qualitative	1.1. Is the qualitative approach appropriate to answer the research question?				
	1.2. Are the qualitative data collection methods adequate to address the research question?				
	1.3. Are the findings adequately derived from the data?				
	1.4. Is the interpretation of results sufficiently substantiated by data?				
	1.5. Is there coherence between qualitative data sources, collection, analysis and interpretation?				
2. Quantitative randomized controlled trials	2.1. Is randomization appropriately performed?				
	2.2. Are the groups comparable at baseline?				
	2.3. Are there complete outcome data?				
	2.4. Are outcome assessors blinded to the intervention provided?				
	2.5 Did the participants adhere to the assigned intervention?				
3. Quantitative non-randomized	3.1. Are the participants representative of the target population?				
	3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)?				
	3.3. Are there complete outcome data?				
	3.4. Are the confounders accounted for in the design and analysis?				
	3.5. During the study period, is the intervention administered (or exposure occurred) as intended?				
4. Quantitative descriptive	4.1. Is the sampling strategy relevant to address the research question?				
	4.2. Is the sample representative of the target population?				
	4.3. Are the measurements appropriate?				
	4.4. Is the risk of nonresponse bias low?				
	4.5. Is the statistical analysis appropriate to answer the research question?				
5. Mixed methods	5.1. Is there an adequate rationale for using a mixed methods design to address the research question?				
	5.2. Are the different components of the study effectively integrated to answer the research question?				
	5.3. Are the outputs of the integration of qualitative and quantitative components adequately interpreted?				
	5.4. Are divergences and inconsistencies between quantitative and qualitative results adequately addressed?				
	5.5. Do the different components of the study adhere to the quality criteria of each tradition of the methods involved?				

Part II: Explanations

3. Quantitative non-randomized studies	Methodological quality criteria
Non-randomized studies are defined as any quantitative studies estimating the effectiveness of an intervention or studying other exposures that do not use randomization to allocate units to comparison groups (Higgins and Green, 2008). Common designs include (this list if not exhaustive): **Non-randomized controlled trials** The intervention is assigned by researchers, but there is no randomization, e.g., a pseudo-randomization. A non-random method of allocation is not reliable in producing alone similar groups. **Cohort study** Subsets of a defined population are assessed as exposed, not exposed, or exposed at different degrees to factors of interest. Participants are followed over time to determine if an outcome occurs (prospective longitudinal). **Case-control study** Cases, e.g., patients, associated with a certain outcome are selected, alongside a corresponding group of controls. Data is collected on whether cases and controls were exposed to the factor under study (retrospective). **Cross-sectional analytic study** At one particular time, the relationship between health-related characteristics (outcome) and other factors (intervention/exposure) is examined. E.g., the frequency of outcomes is compared in different population subgroups according to the presence/absence (or level) of the intervention/exposure. Key references for non-randomized studies: Higgins and Green (2008); Porta et al. (2014); Sterne et al. (2016); Wells et al. (2000)	3.1. Are the participants representative of the target population? Explanations Indicators of representativeness include: clear description of the target population and of the sample (inclusion and exclusion criteria), reasons why certain eligible individuals chose not to participate, and any attempts to achieve a sample of participants that represents the target population.
	3.2. Are measurements appropriate regarding both the outcome and intervention (or exposure)? Explanations Indicators of appropriate measurements include: the variables are clearly defined and accurately measured; the measurements are justified and appropriate for answering the research question; the measurements reflect what they are supposed to measure; validated and reliability tested measures of the intervention/exposure and outcome of interest are used, or variables are measured using 'gold standard'.
	3.3. Are there complete outcome data? Explanations Almost all the participants contributed to almost all measures. There is no absolute and standard cut-off value for acceptable complete outcome data. Agree among your team what is considered complete outcome data in your field (and based on the targeted journal) and apply this uniformly across all the included studies. For example, in the literature, acceptable complete data value ranged from 80% (Thomas et al., 2004; Zaza et al., 2000) to 95% (Higgins et al., 2016). Similarly, different acceptable withdrawal/dropouts rates have been suggested: 5% (de Vet et al., 1997; MacLehose et al., 2000), 20% (Sindhu et al., 1997; Van Tulder et al., 2003) and 30% for follow-up of more than one year (Viswanathan and Berkman, 2012).
	3.4. Are the confounders accounted for in the design and analysis? Explanations Confounders are factors that predict both the outcome of interest and the intervention received/exposure at baseline. They can distort the interpretation of findings and need to be considered in the design and analysis of a non-randomized study. Confounding bias is low if there is no confounding expected, or appropriate methods to control for confounders are used (such as stratification, regression, matching, standardization, and inverse probability weighting).
	3.5 During the study period, is the intervention administered (or exposure occurred) as intended? Explanations For intervention studies, consider whether the participants were treated in a way that is consistent with the planned intervention. Since the intervention is assigned by researchers, consider whether there was a presence of contamination (e.g., the control group may be indirectly exposed to the intervention) or whether unplanned co-interventions were present in one group (Sterne et al., 2016). For observational studies, consider whether changes occurred in the exposure status among the participants. If yes, check if these changes are likely to influence the outcome of interest, were adjusted for, or whether unplanned co-exposures were present in one group (Morgan et al., 2017).

Algorithm for selecting the study categories to rate in the MMAT*



*Adapted from National Institute for Health Care Excellence. (2012). *Methods for the development of nice public health guidance*. London: National Institute for Health and Care Excellence; and Scottish Intercollegiate Guidelines Network. (2017). *Algorithm for classifying study design for questions of effectiveness*. Retrieved December 1, 2017, from http://www.sign.ac.uk/assets/study_design.pdf.

Appendix D

Summary of Risk of Bias Assessments: MMAT (Hong et al., 2018)

Study	Are there clear research questions?	Do the collected data allow to address the research questions?	Are the participants representative of the target population?	Are measurements appropriate regarding both the outcome and intervention (or exposure)?	Are there complete outcome data?	Are the confounders accounted for in the design and analysis?	During the study period, is the intervention administered (or exposure occurred) as intended?
Agha et al. (2024)	Low	Low	Unclear	Low	Low	Low	Low
Akinleye et al. (2024)	Low	Low	Unclear	Low	High	Low	Low
Bourne et al. (2019)	Low	Low	Low	Low	High	Low	Low
Carthon et al. (2024)	Low	Low	Unclear	Low	Unclear	Low	Low
Coombs et al. (2020)	Low	Low	Unclear	Low	High	High	Low
Douglas et al. (2021)	Low	Low	Low	Low	Unclear	Low	Low
Dyrbye et al. (2022)	Low	Low	Low	Low	High	Low	Low
Garcia et al. (2020)	Low	Low	Unclear	Low	High	Low	Low
Khan et al. (2021)	Low	Low	Unclear	Low	High	Low	Low
Lee et al. (2021)	Low	Low	Low	Low	High	Low	Low
Lu et al. (2023)	Low	Low	Unclear	Unclear	High	Low	Low
Martinez et al. (2022)	Low	Low	Unclear	Low	High	Unclear	Low
Merces et al. (2020)	Low	Low	Unclear	Low	High	Unclear	Low
Murray and Chiotu (2024)	Low	Low	Unclear	Low	Low	Low	Low
Nituica et al. (2021)	Low	Low	Low	Low	Low	Low	Low
Parker et al. (2025)	Low	Low	Low	Low	High	Low	Low
Pillado et al. (2022)	Low	Low	Unclear	Unclear	High	High	Low
Robinson et al. (2024)	Low	Low	Unclear	Low	High	Low	Unclear
Sudol et al. (2021)	Low	Low	Low	Low	Low	Low	Low
Taylor et al. (2024)	Low	Low	Unclear	Low	High	Low	Low

Note. Low = Yes; High = No; Unclear = Can't Tell.

Chapter Two: Empirical Study

Identifying Occupational Burnout Subtypes in Healthcare Professionals Accessing the Mind Management Skills for Life Programme Intervention

Abstract

Objective: Burnout is a public health issue among healthcare professionals. Largely, burnout interventions are promising; however, research suggests variability at an individual level. This study aimed to determine the existence of burnout subtypes and examine treatment responses to the Mind Management Skills for Life (MMSFL) programme.

Methods: Secondary analysis was performed using data from two randomised controlled trials testing the effectiveness of MMSFL against a waitlist control group: Chimp Paradox Model Trial 1 (n = 173 nurses) and Trial 2 (n = 294 doctors). The present study used baseline item-level data from the Oldenburg Burnout Inventory and applied machine learning methods (Self-Organizing Maps) to identify clusters of burnout subtypes. Twelve identified subtypes were analysed in regression models to examine whether the subtypes predicted differential treatment responses to MMSFL.

Results: The twelve subtypes were not associated with burnout outcomes after adjusting for baseline severity. Baseline severity did not consistently predict outcomes; however, descriptive analyses indicated a positive relationship between more severe burnout at the start and end of MMSFL across samples. This trend was strongest among doctors from an underrepresented racial-ethnic identity and/or junior roles.

Conclusions: Based on the variables and populations examined in this study, it cannot be concluded how to personalise treatment for MMSFL.

Keywords: *Burnout, mind management, healthcare professionals, machine learning*

Practitioner Points

- Machine learning identified twelve different burnout subtypes.
- External cross-validation is critical to examine the clinical utility of subtypes.
- Future research is needed with larger, more diverse samples, to assess the robustness of preliminary findings.

Introduction

The National Health Service (NHS) is experiencing critical pressures, with chronic levels of occupational burnout among healthcare professionals (HCPs; Nash, 2025). Prolonged workforce shortages, rising patient complexity, and an under-resourced system are key influences on NHS employee burnout (General Medical Council [GMC], 2023). The International Classification of Diseases (ICD-11) defines burnout as an occupational phenomenon, recognising it as a serious health concern associated with the work environment (World Health Organization, 2019). Global prevalence of burnout among medical and nursing roles is reported between 11% and 30% (De Hert, 2020; Woo et al., 2020), increasing up to 52% during the COVID-19 pandemic (Ghahramani et al., 2021). Acute burnout rates in mental health workers have been identified in meta-analytic research (O'Connor et al., 2018), alongside highest ever reported levels of burnout risk in 23% of NHS trainee doctors and 52% of their trainers (GMC, 2023). These alarming statistics reflect the escalating workload pressures and emotionally demanding nature of such roles (Laker et al., 2023), further highlighting burnout as problematic among HCPs.

Burnout in HCPs has been associated with higher levels of stress, depression, and anxiety (Medisauskaite et al., 2023; Salvagioni et al., 2017), leading to decreased job satisfaction (Hall et al., 2020; Taris, 2006) and increased absenteeism (Lee et al., 2011). Critically, burnout in doctors has been linked to increased suicide risk (Harvey et al., 2021). Poorer patient psychotherapy outcomes and increasing medical errors (Delgadillo et al., 2018; Li et al., 2024) are also reported in the burnout literature, posing a substantial threat to the financial status and sustainability of the NHS. Work-related stress absences account for around 40% of HCPs annual sickness and a recent 25% increase in nurses leaving the NHS has been linked to such pressures (Holmes, 2022; Ravalier et al., 2020). The NHS Long Term Workforce Plan (NHS England, 2023) positions staff wellbeing and retention as central priorities, particularly in relation to health and work-life balance; however, the absence of detailed implementation strategies emphasises the importance of further research into effective approaches for mitigating burnout.

Burnout interventions are categorised as individual (e.g., stress management, mindfulness, cognitive behavioural approaches) and organisational (e.g., workload adjustment, peer support, job redesign) approaches (Walsh et al., 2019). Meta-analytic research indicates promising burnout outcomes for both, although organisational interventions show mixed effects as they often target only one domain of burnout (Lee et al., 2016; Panagioti et al., 2017; West et al., 2016). Cognitive behavioural and mindfulness interventions that address personal coping strategies are the most well-established, demonstrating improvements in burnout among HCPs in a Cochrane review and another systematic review (Cohen et al., 2023; Tamminga et al., 2023). Cognitive behavioural techniques show moderate effects for burnout in medical doctors (Clough et al., 2017), as well as 1-year maintenance gains in nurses (Lee et al., 2016). However, the majority of experimental studies report small to moderate effect sizes alongside short-term gains, highlighting how effects may diminish and the need for longer follow-up studies to examine the benefits of these interventions over time.

Digital healthcare interventions (DHIs) offering individual approaches (e.g., web-based interventions, mobile applications, digital support platforms) are increasingly being used to mitigate burnout among HCPs, demonstrating some acceptability and flexibility in the literature (Aye et al., 2024; Dincer & Inangil, 2021). However, multiple studies have found no notable differences between groups when comparing DHIs to established intervention methods or when compared to control groups receiving no treatment (Barrett & Stewart, 2021; Ilola at al., 2024). While DHIs may offer cost-effective solutions, studies have reported barriers including lack of personal relevance and cultural acceptance issues among

HCPs (Zhang et al., 2025). Other workplace factors such as mental health stigma, lack of leadership support, and time constraints have also been identified as challenges to the accessibility of mental health interventions for staff, with the complexity of DHIs adding to some of these pressures (Aye et al., 2024; Paterson et al., 2024).

Research suggests organisations primarily focus on the consequences of burnout, limiting opportunities to address individual factors. In addition, generic interventions that do not meet the unique needs of HCPs often have low engagement (Moe-Byrne et al., 2022; Zhang et al., 2025). These factors heighten burnout vulnerability and create obstacles that limit the effectiveness of support strategies, highlighting the need for interventions to consider demographics, occupational factors, and symptom profiles (Aye et al., 2024; Panagioti et al., 2017). Recent clinical guidelines recommend further research to understand the specific needs of HCPs, considering the effectiveness of individual-level interventions across groups (National Institute for Health and Care Excellence [NICE], 2022). Given the significant impacts of burnout among HCPs, it is critical for research to focus on recognition, prevention, and intervention strategies. This will broaden the understanding of which approaches are effective for different groups, allowing for more targeted, engaging, and accessible support for those most likely to benefit.

The Mind Management Skills for Life (MMSFL) programme is a psychological skills-based intervention developed by Professor Steve Peters, grounded in the Chimp Model of mind management (Peters, 2012). The theory-driven structured programme is evidence-based and designed for individuals and organisations, aiming to reduce burnout and improve wellbeing in HCPs by helping them develop greater emotional self-awareness, regulation strategies, and psychological resilience. Several small pilot studies (unpublished) with NHS staff across primary and secondary care settings suggest that the MMSFL programme may support wellbeing and quality of life. Laker et al. (2023) undertook the first randomised controlled trial (RCT) to evaluate this novel intervention for burnout within the NHS in England, United Kingdom (UK). Published findings indicate significant reductions in burnout among registered mental health nurses with moderate effects, alongside evidence of sustained benefits at six-month follow-up.

To expand on the findings by Laker et al. (2023), a second trial following the same research design and methodology investigated the efficacy of the MMSFL programme with registered doctors in the NHS in England between 2023 and 2025. The study protocol indicates results are yet to be published (https://doi.org/10.1186/ISRCTN14947225). While Laker et al. (2023) provide preliminary support for the efficacy of the MMSFL programme, the evidence remains at an early stage. There is limited understanding regarding which HCPs are most likely to benefit from this intervention and whether its effectiveness varies across burnout subtypes. Burnout is widely recognised as a multidomain construct characterised by emotional exhaustion, depersonalisation, and reduced personal accomplishment (Maslach & Leiter, 2016). However, some research suggests weak correlations between personal accomplishment and burnout outcomes (West et al., 2018). While well-established burnout measures are structured around core domains, such as the exhaustion and disengagement subscales on the Oldenburg Burnout Inventory (OLBI, Demerouti et al., 2001), it is plausible that there might be more diverse subtypes of burnout.

A stronger rationale for examining burnout subtypes can be drawn from the depression literature, where multiple subtypes have been proposed based on clinical presentation and treatment responses (Fava et al., 1997; Lamers et al., 2010; Musil et al., 2018). Recent data-driven studies suggest that depressive subtypes often align with, and may extend beyond, traditional subtypes (Simmonds-Buckley et al., 2021). For example, Núñez et al. (2024) highlighted different depression symptom networks and connectivity patterns in common profiles identified using the Patient Health Questionnaire (PHQ-9). Optimal treatment matching approaches based on individual characteristics have also demonstrated significantly better improvements in several depression studies (Cohen et al., 2022; Huibers et al., 2016; Kappelmann et al., 2020). These findings suggest clustering baseline symptom profiles for burnout may facilitate the identification of clinically relevant subtypes to help predict outcomes and explain differential treatment responses.

Based on the current evidence, burnout appears to be a multifaceted construct that could be influenced by individual, professional, and organisational factors (Maslach & Leiter, 2016). Depression research shows that identifying symptom-level subtypes can achieve optimal treatment outcomes. However, the literature shows burnout interventions are often moderately effectively and inconsistent, with emerging evidence for novel interventions remaining limited (Laker et al., 2023; Lee et al., 2016; Tamminga et al., 2023; West et al., 2016). Understanding burnout subtypes that may differ in their risk factors, severity, and treatment responses has the potential to enable more targeted approaches for specific populations (Beacham et al., 2023). To address such a gap in knowledge, this study aimed to identify burnout subtypes using a validated burnout measure and examine whether subtypes predict differential responses to the innovative MMSFL programme.

Aims, Objectives, and Hypotheses

The overall aim of this study was to determine the existence and clinical significance of burnout subtypes among HCPs. The primary objective was to identify burnout subtypes using the 16-item Oldenburg Burnout Inventory (OLBI). The secondary objective was to examine if HCPs with different subtypes respond differently to the MMSFL programme. This study adopted an inductive data-driven approach and as such, no prior assumptions were made about the specific number or features of subtypes. Based on the two domains proposed on the OLBI (exhaustion and disengagement), it was hypothesised that more than two subtypes would be identified, with no upper limit on the number of subtypes specified a priori. Secondly, it was anticipated that HCPs with different burnout subtypes would show differential treatment responses to the MMSFL programme, such that some would have better outcomes after controlling for baseline severity of burnout.

Ethical Approval

Ethical approval was sought and granted by the University of Sheffield Ethics Committee (see Appendix A) for the secondary analysis of data from the clinical trials. Additional ethics and study information for the two trials is provided in the international register of controlled trials: Trial 1 (https://doi.org/10.1186/ISRCTN34503872) and Trial 2 (https://doi.org/10.1186/ISRCTN14947225).

Methods

Design and Setting

This study is reported in line with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis – Artificial Intelligence (TRIPOD-AI) reporting guidelines for machine learning studies (Collins et al., 2024). The full TRIPOD-AI checklist can be reviewed in Appendix B. This was a secondary data analysis study using pre-existing data collected from HCPs who accessed the MMSFL programme in NHS services in the North of England between 2020 and 2025. The data was provided in two fully anonymised clinical datasets comprising participant outcomes from two independent stepped-wedge RCTs undertaken in several regions across South Yorkshire and the Humber. The first trial was conducted with nurses between 2020 and 2021, named Chimp Paradox Model (CPM) Trial 1 (Laker et al., 2023). The second trial was conducted with doctors between 2023 and 2025, named CMP Trial 2 (CPM Trial Research Group, 2023). The trials will be referred to as CMP 1 and CPM 2 from this point forward.

Participating NHS trusts were the first healthcare services to implement the MMSFL programme with cohorts of HCPs, measuring the collective impact of the group-based intervention on levels of burnout. In the trials, the programme was delivered by psychological skills mentors across eight digital sessions. Participants were not blinded due to the nature of the intervention. Intention-to-treat principles were utilised, comparing mean burnout outcomes between group 1 (immediate intervention) and group 2 (waitlist control) across four time points: (a) baseline; (b) after group 1 completed MMSFL; (c) after group 2 completed MMSFL; and (d) six-months follow-up. The Health Research Authority approved the collation of the datasets including participant outcomes for research purposes. The present study was pre-registered via the Open Science Framework (OSF) platform: https://osf.io/j5ukb/?view_only=6249145107174d2caf89ef5467150b68

Inclusion and Exclusion Criteria

Two clinical trial datasets were obtained from CPM 1 and CPM 2. Eligibility criteria is presented in Table 1. If participants met these criteria, they were invited to take part in the study and randomly assigned to the immediate intervention or waitlist control group using a computerised randomisation schedule. Recruitment was carried out in medical hospital and community-based NHS services for both trials, and promoted through several communication channels (e.g., communal posters, electronic newsletters, staff intranet). CPM 1 recruitment was undertaken in February 2020, and CPM 2 between September and November 2023. Informed consent processes were fulfilled and involvement in the trials was voluntary.

Table 1

Inclusion and Exclusion Criteria for CPM Trials

	Inclusion		**Exclusion**	
	CPM 1	**CPM 2**	**CPM 1**	**CPM 2**
Age	Aged 18 or over.	Aged 18 or over.	Aged 18 or under.	Aged 18 or under.
Employment	Full or part-time NHS employment.	Full or part-time NHS employment.	Nurses not in clinical service (e.g., sick leave,	Doctors not in clinical service at (e.g., sick leave,

			maternity, suspended).	maternity, suspended).
Professional Registration	Registration with the NMC.	Registered with the GMC.	No active professional registration.	No active professional registration.
NHS Setting and Role	Clinical role in NHS community nursing service (RDaSH).	Clinical role in any specialty in NHS medical hospital and community services.	Services other than home/clinic-based care for children and adults.	Managerial, supervisory, or educational role.
Therapy Input	Not accessing talking therapies for mental health difficulties.	Not accessing talking therapies for mental health difficulties.	Referred/accessing any concurrent psychological intervention.	Referred/accessing any concurrent psychological intervention.

Note. CMP 1 = Chimp Paradox Model Trial 1; CPM 2 = Chimp Paradox Model Trial 2; NHS = National Health Service; NMC = Nursing and Midwifery Council; GMC = General Medical Council; RDaSH = Rotherham Doncaster & South Humber NHS Foundation Trust.

Intervention

The MMSFL programme was delivered to all participants recruited in the CPM trials through a structured and group-based self-help intervention across eight weeks (one session per week). The sessions were offered via a digital platform with up to 30 participants in each group, comprising a didactic presentation, group-based discussions, self-reflection, and written exercises. The programme provides an understanding of how the mind is organised and how the emotional system operates, incorporating neuroscience and psychological theory. The workshops focus learning around the following areas: improved relationships with self and others, communication, impact of environmental factors, health and wellbeing, and personal development. The psychological skills mentors facilitated group-based guided exercises for skill development, including cognitive and mindfulness-based strategies. Supplementary written materials were provided to help consolidate learning and to support the application of skills in daily work and personal situations. The course handbook provides more information on the workshops (Peters, 2012).

Outcome Measure

The Oldenburg Burnout Inventory (OLBI; Demerouti et al., 2001)

The OLBI (see Appendix C) was the primary measure used in the CPM trials, a 16-item questionnaire designed to measure two domains of burnout: emotional exhaustion (OLBI-E; 8 items) and disengagement (OLBI-D; 8 items). Exhaustion refers to physical, emotional, and cognition depletion in relation to job demands. Disengagement reflects the psychological detachment (e.g., feelings of cynicism) from work. Items are scored on a four-point Likert scale between 1 (strongly agree) and 4 (strongly disagree) with some items reverse-scored. Total scores range from 16 to 64 with higher scores indicative of increased burnout. Leclercq et al. (2021) defined burnout severity across three categories: mild burnout (43 and below), moderate burnout (between 44 and 51), and severe burnout (53 and above). Good internal consistency (α = .74 to .87) is indicated for the OLBI in both CPM trials along with its suitability for use in occupational populations (CPM Trial Research Group, 2023; Laker et al., 2023), aligning with the psychometric properties reported in the wider burnout literature (Halbesleben & Demerouti, 2005). In the present study, item-level data on the OLBI at baseline and post-treatment was used to identify the burnout subtypes.

Sample Characteristics

The study samples were independent of each other, including registered nurses (n = 173) from CPM 1 and registered doctors (n = 294) from CPM 2. The majority of participants were female (CPM 1, 75.1%; CPM 2, 66.7%) and the average age ranged from 40 to 45 years across samples. In terms of racial-ethnic identity, 1.7% of CPM 1 participants identified from an underrepresented group compared to around a third of CPM 2 participants. Samples had comparable proportions of White British participants. The mean number of attended sessions on the MMSFL programme in CPM 1 was 4.08 (3.19). Training grade was disaggregated in the CPM 2 sample, with 48.6% consultants, 43.5% junior doctors, and 7.8% preferring not to disclose their role. Full sample characteristics are summarised in Table 2.

Table 2

Participant Demographic and Clinical Characteristics

	Training Sample	Test Sample
Participant Characteristics	***n* = 173**	***n* = 294**
Demographic Characteristics	**Mean (SD) / %**	**Mean (SD) / %**
Gender %		
Female	75.1	66.7
Male	6.9	33.3
Unknown	17.9	0
Age[a]	45.14 (9.40)	40.03 (9.22)
Employment %		
Full Time	61.8	48.2
Part Time	21.5	51.5
Unknown	16.8	0.3
Ethnicity %		
White British	76.3	61.8
Underrepresented	1.7	36.8
Unknown	21.8	1.7
Training Grade		
Nurse	100	0
Junior Doctor	0	43.4
Consultant	0	48.6
Unknown	0	7.8
Clinical Characteristics		

	Baseline OLBI[a]	40.38 (5.63)	42.80 (5.94)
	Post-Treatment OLBI[a]	35.77 (6.03)	39.40 (6.32)
	Sessions Attended[a]	4.08 (3.19)	-

Note. SD = Standard Deviation; OLBI = Oldenburg Burnout Inventory; Underrepresented Racial-Ethnic Identity = Asian (16.4%), Black (3.4%), Chinese (1.4%), multiracial (4.4%), Other (2.7%); Junior Doctor = Foundation (4.4%), Core Training (13.9%), Higher Training (23.1%), Vocational Training Scheme (2.0%).

[a] Means (SD) reported for demographic characteristics.

Sample Size Calculation

There is no precedent study investigating burnout subtypes in this setting or prior information about the specific design factors required. Methodological and simulation studies were therefore reviewed to determine the minimal necessary sample to identify subgroups of cases using Self-Organizing Maps (SOM), a machine learning technique that learns to cluster data into subgroups based on similarity (Kohonen, 2013). A methodological study by Kiang et al. (2007) found the SOM method outperformed alternative clustering methods such as K-means cluster analysis and factor analysis, even at relatively small sample sizes. In this simulation study, the performance of the SOM method was found to be insensitive to larger sample sizes above the minimum threshold of $n = 50$. Classification accuracy remained stable when compared to increasingly large sample sizes up to $n = 1600$. SOM consistently displayed higher accuracy at segmenting distinctive clusters relative to factor analytic methods in all comparisons of various sample sizes. Moreover, following the guidelines by Kiang et al. (2007), a minimum sample of $n = 50$ cases was proposed in the present study samples for the clustering and cross-validation procedures, which was achieved.

Generalized Linear Models (GLM) were applied to examine predictors of treatment outcomes. Assuming the dependent variable can be predicted with a moderate effect size (d = 0.50), with 80% power and using an alpha level of $p = 0.05$, it was estimated that a minimum of $n = 76$ cases were required in each of the samples (i.e., CPM1 and CPM 2) to perform a multiple regression analysis according to Cohen's rule (Cohen, 1992). There is no precedent study using the SOM clustering technique as a form of data analysis to predict burnout outcomes, and as such there is no prior evidence that can be used to inform a sample size calculation for this step. Adopting a conservative approach, and following guidance by Cohen (1992), it was assumed that treatment response could be predicted with a moderate effect size, rather than a large effect size, which is typical of psychotherapy outcome prediction studies (Lutz et al., 2021).

Data Preparation

The datasets were obtained from two independent groups of HCPs, meaning they could be identified by professional role. To enable the external cross-validation procedure, the datasets were classified as a training sample (CPM 1; $n = 173$ nurses) and a statistically independent test sample (CPM 2; $n = 294$ doctors). Identifying the datasets this way allowed the cross-validation procedure to be undertaken across clinical professions, assessing real-world applicability of the clustering model for different professionals (Ramspek et al., 2020). While the data could have been partitioned using a within-profession split for each dataset, such as a 50:50 split of nurses for the training and test sample, this could have increased the risk of the clusters reflecting the profession-specific context (e.g., occupational factors) rather than assessing the accuracy and generalisability of the burnout-related subtypes across different professions (Delgadillo, 2021).

Missing data was previously imputed in the original trials using an expectation-maximization algorithm, prior to data analysis. In the present study, categorical variables were transformed into binary variables for the purpose of analysis if these were available (e.g., racial-ethnic identity became White British or underrepresented). Participants were required to have completed all item-level data (i.e., 16 items) on the OLBI at the first session (i.e., baseline) of the MMSFL programme to enable the development of burnout subtypes. Datasets were stored on a secure university network drive with restricted researcher access in line with the University of Sheffield data protection policies. Confidentiality was protected at all stages of data collection for both CPM trials using unique case pseudonyms.

Data Analysis

Four stages of data analysis were completed to achieve the study objectives. First, symptom-specific burnout subtypes were identified in the training sample using baseline item-level OLBI indicators. Next, the subtypes were examined for differential treatment responses to the MMSFL programme. Then, the trained clustering algorithm used to identify subtypes in the training sample was applied to the independent test sample as part of the external cross-validation procedure. Finally, treatment responses from the training sample were tested for replication in the test sample using the subtypes identified during cross-validation along with sensitivity analyses. Data analysis was undertaken on a restricted-access drive. The datasets will be held at the University of Sheffield for a minimum of 5-10 years after conclusion of the clinical trials.

Cluster Identification Using Self-Organizing Maps

A clustering algorithm was developed in the training sample to identify subtypes of burnout. This was undertaken using SOM, an unsupervised machine learning method where neural networks are used to form clusters that represent distinct subgroups with similar characteristics, without predefining the number or features of clusters (Galkin et al., 2022; Kohonen, 2013). During the learning process of the clustering algorithm, the nodes in the network adjust their weight according to the lateral feedback, further grouping together in clusters to represent collections of data with related properties (Kiang, 2001). The IBM SPSS Modeler software facilitated the clustering process using baseline item-level data from the OLBI, of which all participants completed. The identified clusters (i.e., cluster membership) represent burnout subtypes in this study. The perceived accuracy of clusters was assessed using the silhouette measure of cohesion and separation (Yin, 2008), evaluating how similar a case is to its own cluster in comparison to others. A silhouette index was generated, aiding selection of the final SOM model.

Cluster Membership and Treatment Outcomes

A Generalized Linear Model (GLM) was conducted in the training sample to examine the relationship between the identified burnout subtypes and MMSFL outcomes. GLMs are widely used in machine learning studies as they are designed to handle diverse data types beyond a standard linear regression, especially when variables do not adhere to normal distribution assumptions (Mamun & Paul, 2023). The GLM tested whether the burnout subtypes from the initial clustering process had clinical utility and could predict treatment outcomes beyond baseline severity. Burnout subtypes were referred to as clusters after SOM assigned the subtype profiles to clusters ranging from 1 to 12. Regression models were built with post-treatment OLBI scores as the dependent variable and cluster membership as the independent variable, adjusting for baseline severity. The R-squared (R^2) measure of variance (Cohen, 1988) was used as the index of accuracy to understand which clusters might predict post-treatment outcomes with greater accuracy.

External Cross-Validation Procedure

To determine the generalisability and credibility of findings, the SOM clustering model developed in the training sample was applied to the independent test sample to test its performance, following an external cross-validation method (Delgadillo, 2021). This process involved assigning each participant from the test sample to one of the twelve clusters using the SOM clustering algorithm. Next, the GLM regression model that was run in the training sample was conducted in the test sample to examine whether the results replicated in a new clinical sample. Again, post-treatment OLBI scores were inputted as the dependent variable, alongside cluster membership as the independent variable, adjusting for baseline severity. An index of accuracy (R^2) was also calculated for the test sample.

Sensitivity Analyses

Sensitivity analyses were carried out in the test sample to understand the influence of specific variables on treatment outcomes alongside cluster membership. Firstly, a GLM analysed subgroups of participants classified as junior doctor (i.e., foundation, core training, higher training, vocational training scheme) or consultant. It was expected that smaller subgroupings of roles would not be large enough to conduct an adequately powered analysis. Another GLM was subsequently run for participants who identified from either a White British or underrepresented racial-ethnic identity (i.e., Asian, Black, Chinese, multiracial, Other). These analyses helped to examine prediction accuracy (Salciccioli et al., 2016), specifically whether demographic and/or work-related characteristics were associated with post-treatment OLBI scores in the test sample.

Results

Cluster Identification of Burnout Subtypes

Two SOM models were developed in the training sample to determine which would provide the best model fit, using the silhouette measure of cohesion and separation to assess cluster quality. This resulted in an exponential model of 62 clusters and a linear model that provided a maximum limit of 12 clusters. Guided by the silhouette index, the simpler and more interpretable linear model was selected as the exponential model did not improve model fit. The average silhouette index for the clusters was 0.1, suggesting there may be overlap of clusters and closer data points in individual and neighbouring clusters (Bowen & Siegler, 2024). In the training sample (CPM 1 participants), the SOM method identified 12 different burnout subtypes. The overall OLBI baseline mean score was 40.38 (5.63), suggesting mild yet high variability in the spread of burnout levels. Cluster 3 and 10 were the most prevalent subtypes at 14.5% and 12.7%, respectively. Least prevalent clusters were 2 and 11 (see Table 3). Visual bar charts are presented in Figure 1 to show the subtype structures, including item-level differences in mean burnout severity and 95% confidence interval error bars.

Table 3

Frequency and Percentage of Clusters and Mean Total OLBI at Baseline

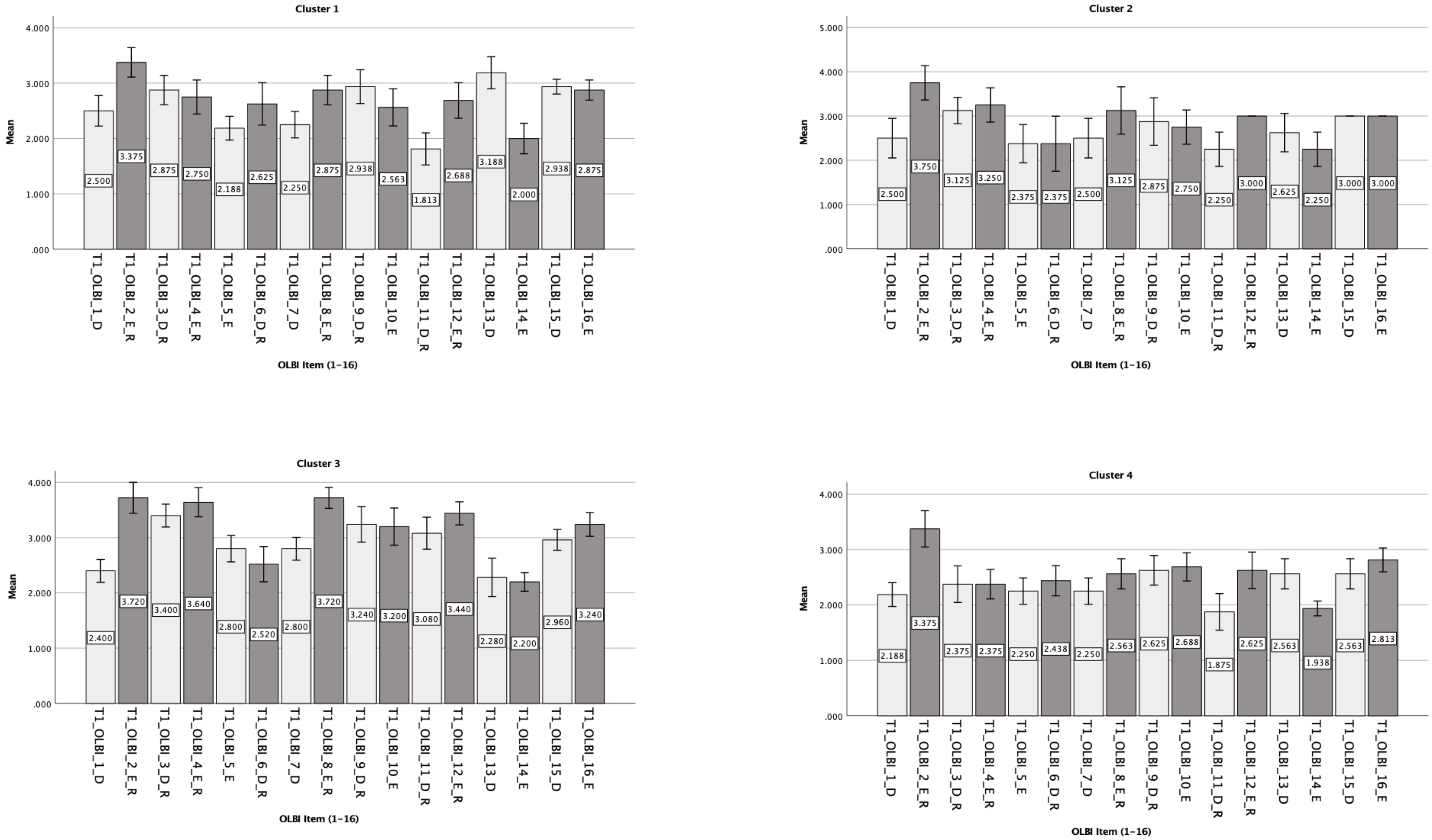
	Training Sample			Test Sample		
Cluster Number	***n***	**%**	**Mean OLBI Score (SD)**	***n***	**%**	**Mean OLBI Score (SD)**
1	16	9.2	42.44 (2.66)	41	13.9	42.61 (1.76)
2	8	4.6	44.75 (1.49)	32	10.9	46.19 (1.55)
3	25	14.5	48.64 (2.93)	69	23.5	50.30 (3.38)
4	16	9.2	39.50 (1.20)	22	7.5	39.00 (1.48)
5	12	6.9	41.83 (1.11)	23	7.8	41.78 (0.95)
6	14	8.1	44.14 (1.46)	18	6.1	44.50 (1.62)
7	15	8.7	36.40 (1.06)	21	7.1	36.33 (1.65)
8	10	5.8	38.10 (0.99)	9	3.1	37.78 (0.83)
9	11	6.4	42.00 (2.14)	16	5.4	40.50 (1.46)

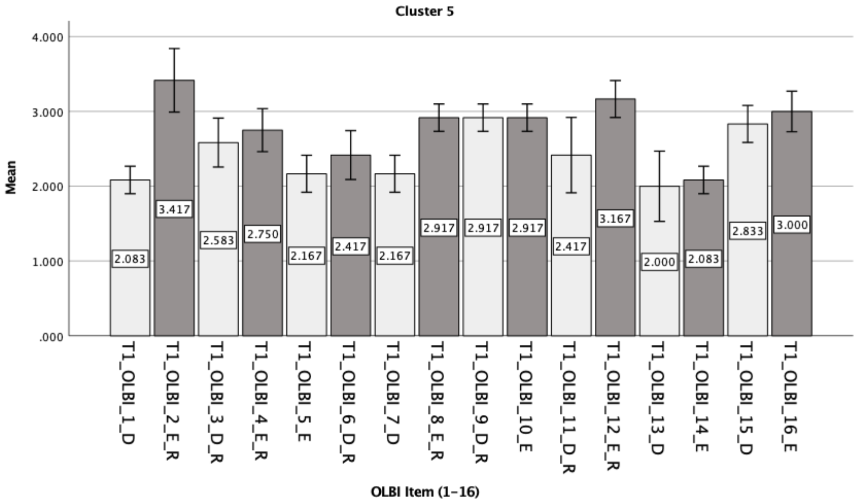
10	22	12.7	31.18 (3.39)	19	6.5	31.68 (3.61)
11	8	4.6	34.75 (1.98)	10	3.4	35.30 (1.89)
12	16	9.2	39.25 (2.59)	14	4.8	40.14 (2.14)
Total	173	100	40.38 (5.63)	294	100	42.80 (5.94)

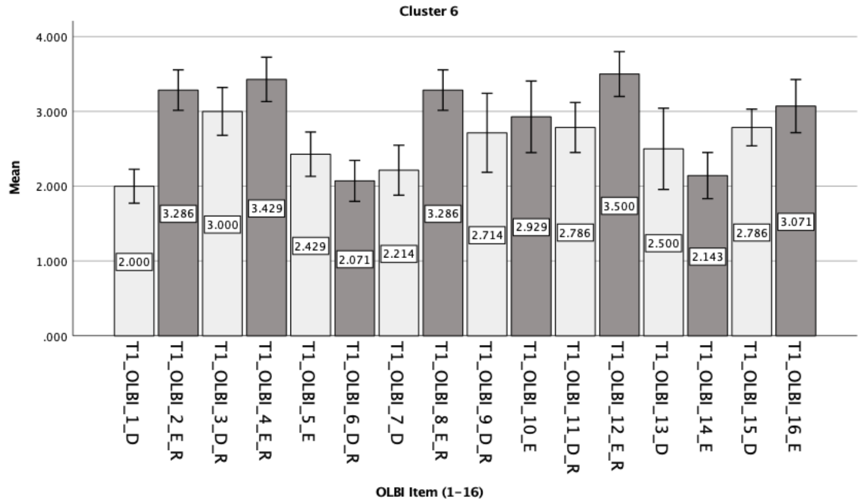
Note. OLBI = Oldenburg Burnout Inventory; SD = Standard Deviation.

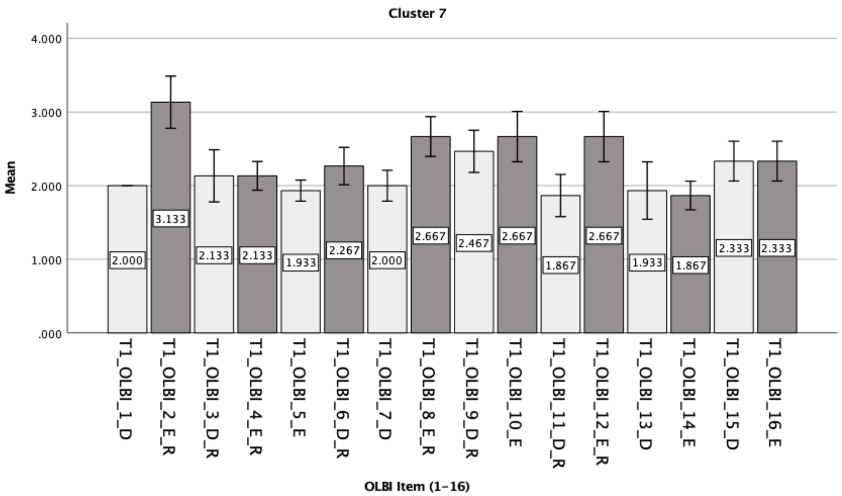
Figure 1

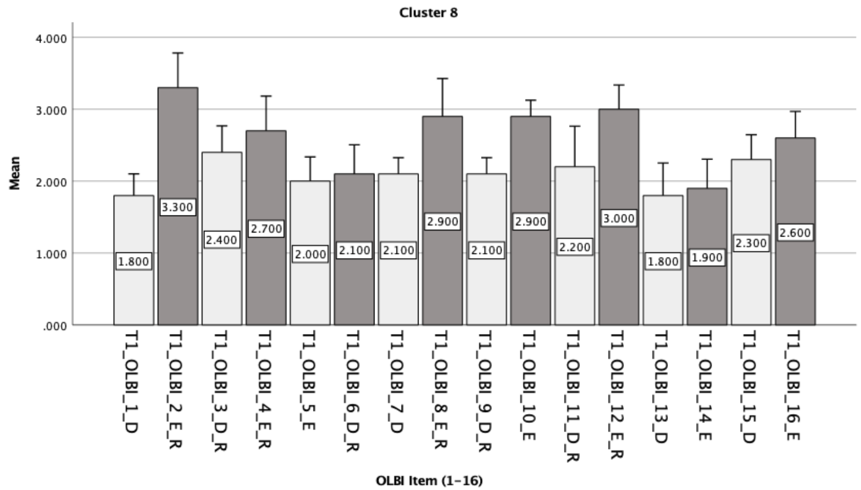
Mean Baseline Score of Each Item on the OLBI for Clusters 1-12 in the Training Sample (CPM 1)

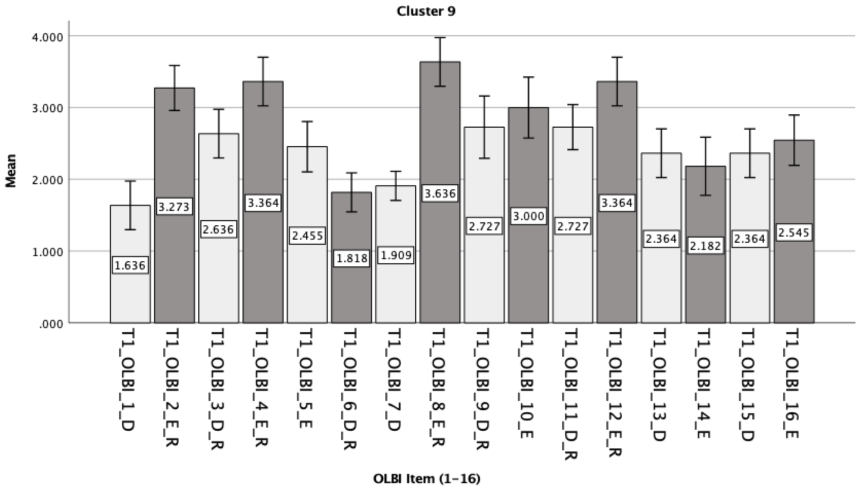


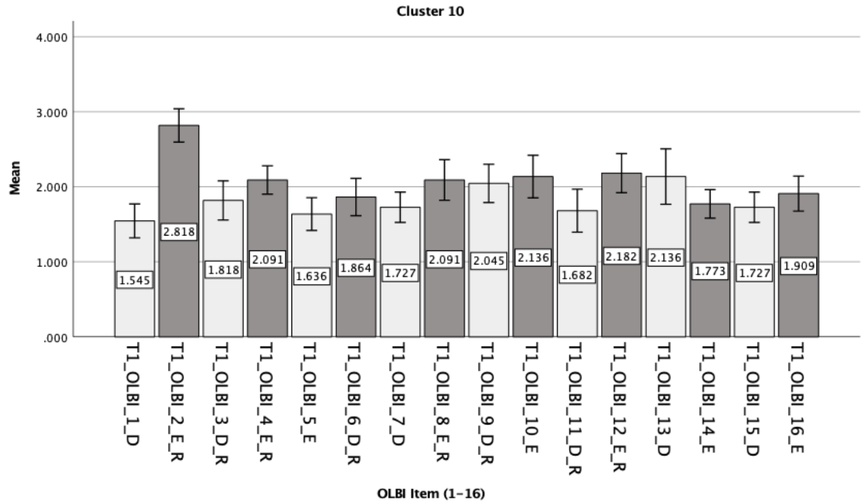

Cluster 5
Mean
OLBI Item (1-16)

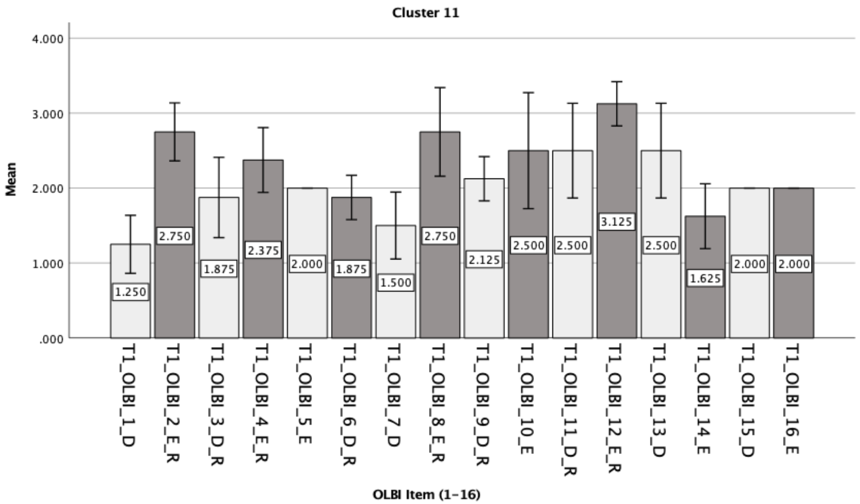

Cluster 6
Mean
OLBI Item (1-16)

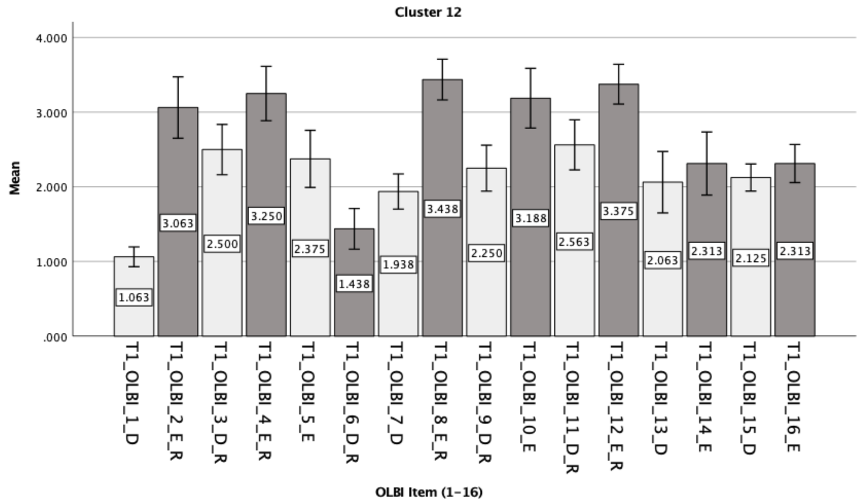

Cluster 7
Mean
OLBI Item (1-16)


Cluster 8
Mean
OLBI Item (1-16)


Cluster 9
Mean
4.000
3.000
2.000
1.000
.000
1.636
3.273
2.636
3.364
2.455
1.818
1.909
3.636
2.727
3.000
2.727
3.364
2.364
2.182
2.364
2.545
T1_OLBI_1_D
T1_OLBI_2_E_R
T1_OLBI_3_D_R
T1_OLBI_4_E_R
T1_OLBI_5_E
T1_OLBI_6_D_R
T1_OLBI_7_D
T1_OLBI_8_E_R
T1_OLBI_9_D_R
T1_OLBI_10_E
T1_OLBI_11_D_R
T1_OLBI_12_E_R
T1_OLBI_13_D
T1_OLBI_14_E
T1_OLBI_15_D
T1_OLBI_16_E
OLBI Item (1-16)


Cluster 10
Mean
4.000
3.000
2.000
1.000
.000
1.545
2.818
1.818
2.091
1.636
1.864
1.727
2.091
2.045
2.136
1.682
2.182
2.136
1.773
1.727
1.909
T1_OLBI_1_D
T1_OLBI_2_E_R
T1_OLBI_3_D_R
T1_OLBI_4_E_R
T1_OLBI_5_E
T1_OLBI_6_D_R
T1_OLBI_7_D
T1_OLBI_8_E_R
T1_OLBI_9_D_R
T1_OLBI_10_E
T1_OLBI_11_D_R
T1_OLBI_12_E_R
T1_OLBI_13_D
T1_OLBI_14_E
T1_OLBI_15_D
T1_OLBI_16_E
OLBI Item (1-16)


Cluster 11
Mean
4.000
3.000
2.000
1.000
.000
1.250
2.750
1.875
2.375
2.000
1.875
1.500
2.750
2.125
2.500
2.500
3.125
2.500
1.625
2.000
2.000
T1_OLBI_1_D
T1_OLBI_2_E_R
T1_OLBI_3_D_R
T1_OLBI_4_E_R
T1_OLBI_5_E
T1_OLBI_6_D_R
T1_OLBI_7_D
T1_OLBI_8_E_R
T1_OLBI_9_D_R
T1_OLBI_10_E
T1_OLBI_11_D_R
T1_OLBI_12_E_R
T1_OLBI_13_D
T1_OLBI_14_E
T1_OLBI_15_D
T1_OLBI_16_E
OLBI Item (1-16)


Cluster 12
Mean
4.000
3.000
2.000
1.000
.000
1.063
3.063
2.500
3.250
2.375
1.438
1.938
3.438
2.250
3.188
2.563
3.375
2.063
2.313
2.125
2.313
T1_OLBI_1_D
T1_OLBI_2_E_R
T1_OLBI_3_D_R
T1_OLBI_4_E_R
T1_OLBI_5_E
T1_OLBI_6_D_R
T1_OLBI_7_D
T1_OLBI_8_E_R
T1_OLBI_9_D_R
T1_OLBI_10_E
T1_OLBI_11_D_R
T1_OLBI_12_E_R
T1_OLBI_13_D
T1_OLBI_14_E
T1_OLBI_15_D
T1_OLBI_16_E
OLBI Item (1-16)

Predictors of Treatment Outcomes

To examine the relationship between identified burnout subtypes and treatment outcomes in the training sample, a GLM was conducted (see Table 4). Initially, the independent variable was included in the model (i.e., SOM cluster membership) as a predictor of post-treatment OLBI scores without controlling for other covariates. Cluster 1 was indicated as the reference group. The unadjusted model is displayed in Table 4, indicating clusters 8 (B = -5.26, SE = 2.45, p = .035), 10 (B = -9.36, SE = 2.25, p <.001), and 12 (B = -5.30, SE = 2.37, p = .028) as significant predictors of post-treatment OLBI scores. With an R^2 = .28, cluster membership accounted for 28% of the variance in treatment outcomes. Based on initial examination, clusters mainly showed negative coefficients, suggesting members of these clusters had lower post-treatment OLBI scores, and experienced better outcomes compared to the reference group (i.e., cluster 1).

Table 4

Summary of Unadjusted GLM Analysis for Variables Predicting Treatment Outcomes

	Training Sample			
Cluster Number	**Coefficient *B* (*SE*)**	***t***	***p***	**95% CI**
1	b	b	b	b
2	-.14 (2.68)	-.05	.959	[-5.46, 5.18]
3	2.29 (2.12)	1.08	.284	[-1.93, 6.52]
4	-.64 (2.68)	-.24	.813	[-5.96, 4.68]
5	-2.78 (2.55)	-1.09	.279	[-7.85, 2.29]
6	-3.44 (2.84)	-1.20	.230	[-9.09, 2.22]
7	-4.41 (2.37)	-1.86	.066	[-9.13, .30]
8	-5.26* (2.45)	-2.15	.035	[-10.13, -.39]
9	-2.26 (2.45)	-.92	.359	[-7.13, 2.61]
10	-9.36** (2.25)	-4.17	<.001	[-13.83, -4.90]
11	-4.84 (2.84)	-1.70	.093	[-10.49, .82]

12	-5.30* (2.37)	-2.24	.028	[-10.01, -.59]

Note. SE = Standard Error; *p = < .05; **p = <.001.

[a] Dependent Variable: Post-Treatment OLBI.

[b] Cluster 1 was indicated as the reference group.

In the adjusted model, the explained variance in treatment outcomes for the training sample increased by 2% (adjusted R^2 = .30). No clusters were statistically significant predictors in this model after controlling for baseline OLBI (see Table 5). On visual inspection (see Figure 2) of baseline and post-treatment OLBI scores via a scatterplot, a positive association is observed with several outliers showing more severe pre-treatment and post-treatment scores. This indicates that baseline severity explained a larger amount of variance in post-treatment OLBI, rendering the coefficients for burnout subtypes redundant. In short, several clusters (i.e., clusters 8, 10, and 12) initially appeared to influence treatment outcomes; however, no clusters predicted outcomes after controlling for baseline severity. Moreover, it appears there is no additional predictive value in the subtypes than what is observed at baseline, meaning members of certain clusters might have had better treatment outcomes due to experiencing mild to moderate burnout levels at baseline.

Table 5

Summary of Adjusted GLM Analysis for Variables Predicting Treatment Outcomes

	Training Sample			
Cluster Number	**Coefficient *B* (*SE*)**	***t***	***p***	**95% CI**
1	[b]	[b]	[b]	[b]
2	-1.37 (2.76)	-.50	.621	[-6.85, 4.11]
3	-.35 (2.66)	-.13	.896	[-5.64, 4.95]
4	.54 (2.75)	.20	.844	[-4.93, 6.01]
5	-2.53 (2.53)	-1.00	.321	[-7.56, 2.50]
6	-3.94 (2.84)	-1.39	.168	[-9.58, 1.70]
7	-2.00 (2.78)	-.72	.475	[-7.53, 3.54]

8	-3.64 (2.63)	-1.38	.170	[-8.86, 1.59]
9	-2.45 (2.43)	-1.01	.317	[-7.28, 2.39]
10	-4.29 (3.84)	-1.12	.267	[-11.93, 3.35]
11	-2.20 (3.26)	-.68	.501	[-8.67, 4.27]
12	-4.49 (2.40)	-1.87	.065	[-9.27, .28]
Baseline OLBI	.41[a] (.26)	1.62	.109	[-.09, .92]

Note. SE = Standard Error; **p* = < .05; ***p* = <.001.

[a] Dependent Variable: Post-Treatment OLBI.

[b] Cluster 1 was indicated as the reference group.

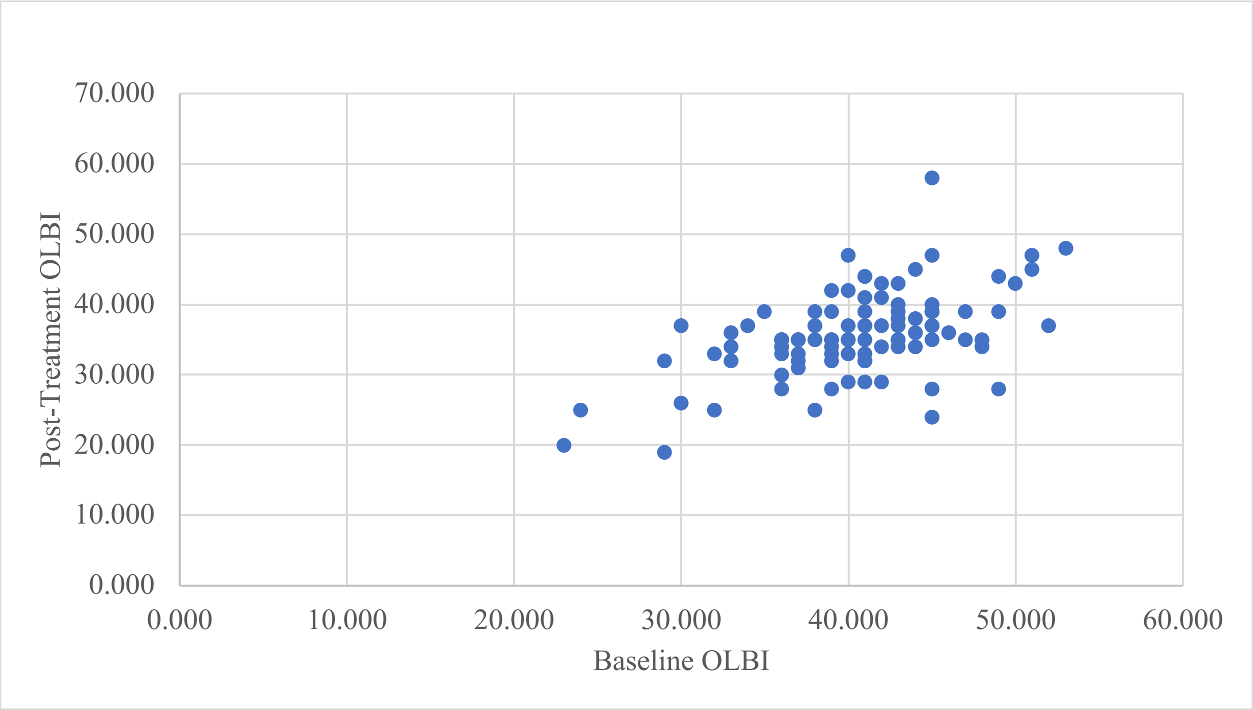
External Cross-Validation of Predictors

The trained clustering algorithm was applied to the test sample (CPM 2 participants) to assess generalisability of subtypes from the training sample as part of the external cross-validation procedure (see Table 3). Across samples, similar frequencies of cases were assigned to each cluster and mean baseline OLBI remained consistent, suggesting increased generalisability of the identified subtypes. To analyse the relationship between identified subtypes and treatment outcomes in the test sample, an identical GLM was repeated (see Table 6). Visual bar charts were generated to show the structure of each of the 12 different subtypes in the test sample (see Appendix D). Comparable to the training sample, the initial model inputted cluster membership as the independent variable without controlling for baseline OLBI. The unadjusted model identified several clusters as predictors of post-treatment OLBI with an R^2 = .42, indicating that cluster membership accounted for 42% of the variance in treatment outcomes. Clusters 2 (B = 3.75, SE = 1.68, p = .027) and 3 (B = 7.26, SE = 1.36, p <.001) displayed positive coefficients, suggesting members of these subtypes had more severe post-treatment OLBI compared to the reference group. Conversely, cluster 10 (B = -6.35, SE = 1.90, p = .001) yielded a negative coefficient, indicating members of this cluster were associated with less severe post-treatment scores.

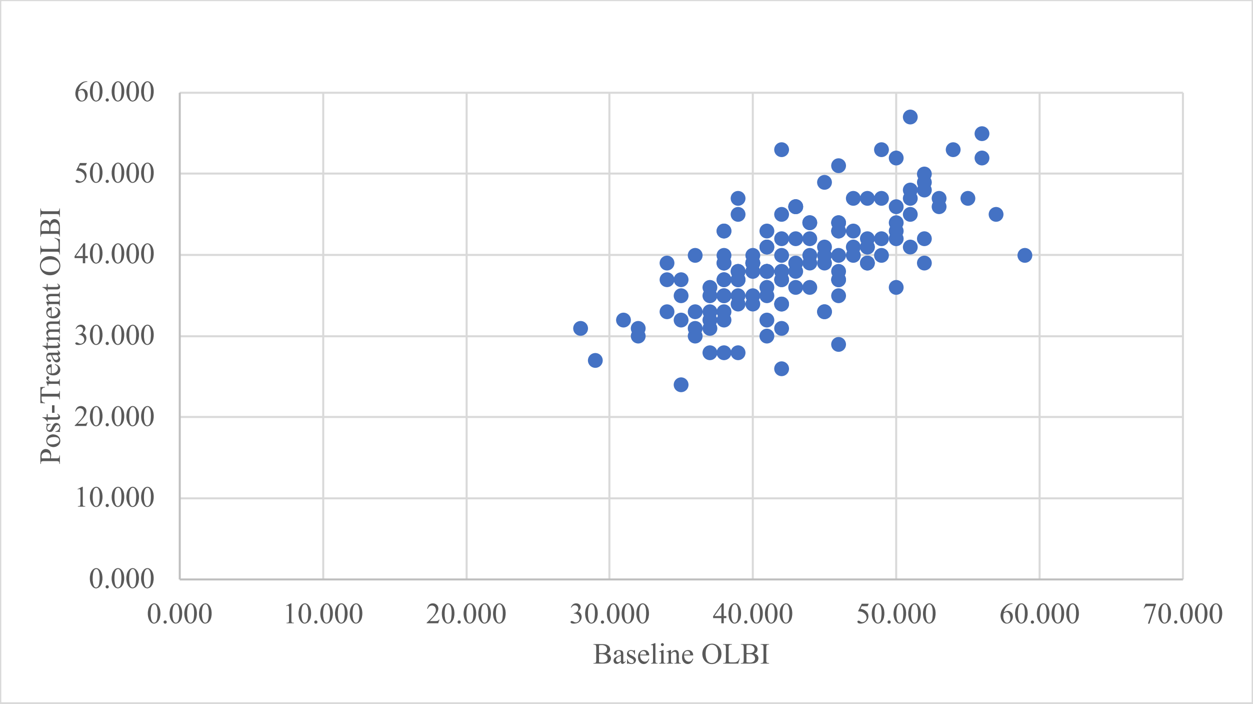
Figure 2

Scatterplots of Baseline and Post-Treatment Outcomes in the Training and Test Samples

Training Sample



Test Sample



Note. The figures show positive relationships between baseline and post-treatment OLBI scores, indicating more severe baseline burnout tends to relate to more severe burnout after MMSFL.

Table 6

Summary of Unadjusted GLM Analysis for Variables Predicting Treatment Outcomes

	Test Sample			
Cluster Number	**Coefficient *B* (*SE*)**	***t***	***p***	**95% CI**
1	[b]	[b]	[b]	[b]
2	3.75* (1.68)	2.23	.027	[.43, 7.07]
3	7.26** (1.36)	5.32	<.001	[4.56, 9.96]
4	-.31 (1.75)	-.18	.859	[-3.78, 3.15]
5	.65 (1.90)	.34	.733	[-3.12, 4.42]
6	1.15 (3.04)	.38	.706	[-4.87, 7.17]
7	-3.49 (1.85)	-1.89	.061	[-7.14, .17]
8	-.85 (2.69)	-.32	.753	[-6.18, 4.48]
9	-.85 (1.90)	-.45	.656	[-4.62, 2.92]
10	-6.35** (1.90)	-3.34	.001	[-10.12, -2.58]
11	-6.35 (3.65)	-1.74	.084	[-13.56, .86]
12	2.29 (2.16)	1.06	.290	[-1.98, 6.56]

Note. SE = Standard Error; **p* = < .05; ***p* = <.001.

[a] Dependent Variable: Post-Treatment OLBI.

[b] Cluster 1 was indicated as the reference group.

In the adjusted model, baseline OLBI scores were added (see Table 7), increasing the explained variance in treatment outcomes for the test sample by 3% (adjusted $R^2 = .45$). No clusters remained significant after controlling for baseline severity, eliminating the initial predictive value of cluster membership. Baseline OLBI ($B = .52$, $SE = .19$, $p = .006$) was statistically significant with a positive coefficient, meaning cases with more severe baseline OLBI experienced more severe symptoms post-treatment. Visualisation (see Figure 2) of baseline and post-treatment OLBI scores confirmed a strong positive association in the test sample. In short, clusters 2, 3, and 10 were initially significant predictors of treatment outcomes. In the adjusted model for the test sample, only baseline OLBI remained a significant predictor; however, this was not a consistent pattern found in the training sample. This indicates that burnout subtypes and baseline severity do not consistently predict outcomes across the samples in the present study.

Table 7

Summary of Adjusted GLM Analysis for Variables Predicting Treatment Outcomes

	Test Sample			
Cluster Number	**Coefficient *B* (*SE*)**	***t***	***p***	**95% CI**
1	[b]	[b]	[b]	[b]
2	1.74 (1.79)	.98	.331	[-1.79, 5.28]
3	3.08 (2.00)	1.54	.126	[-.88, 7.04]
4	1.58 (1.84)	.86	.392	[-2.06, 5.21]
5	1.09 (1.86)	.59	.558	[-2.59, 4.78]
6	.05 (2.99)	.02	.988	[-5.88, 5.97]
7	-.40 (2.11)	-.19	.848	[-4.58, 3.77]
8	1.52 (2.76)	.55	.582	[-3.94, 6.98]
9	.22 (1.90)	.12	.908	[-3.53, 3.97]
10	-1.22 (2.61)	-.47	.642	[-6.39, 3.95]
11	-2.42 (3.82)	-.63	.529	[-9.98, 5.15]
12	3.55 (2.15)	1.65	.102	[-.71, 7.81]
Baseline OLBI	.52[a]* (.19)	2.79	.006	[.15, .89]

Note. SE = Standard Error; **p* = < .05; ***p* = <.001.

[a] Dependent Variable: Post-Treatment OLBI.

[b] Cluster 1 was indicated as the reference group.

Sensitivity Analyses

Finally, a series of sensitivity analyses were performed using GLM, assessing the robustness of results found in the test sample (see Table 8). Post-treatment OLBI scores were inputted as the dependent variable in this model and cases were analysed alongside cluster membership if they identified as (a) junior doctor; (b) consultant; (c) White British; or (d) underrepresented. The model controlled for baseline severity. Beta coefficients that diminished to exactly 0 were indicative of variables that had no prognostic value. Relative to the reference cluster, junior doctors in clusters 4 (B = 6.91, SE = 2.68, p = .013) and 12 (B = 9.91, SE = 3.50, p = .007) experienced more severe post-treatment outcomes after controlling for baseline severity. No significant effects were observed for consultants, or the White British group based on cluster membership. In addition, baseline OLBI was a significant predictor of post-treatment outcomes in the subgroup of junior doctors (B = .87, SE = .30, p = .005) and those identifying from an underrepresented racial-ethnic group (B = .84, SE = .36, p = .025), indicating that more severe baseline severity in these cases was associated with more severe post-treatment burnout.

Table 8

Summary of GLM Analyses for Sample Characteristics in the Test Sample

	Job Role		**Ethnicity**	
	Coefficient *B* (*SE*)		**Coefficient *B* (*SE*)**	
Cluster Number	**Junior Doctors**	**Consultants**	**White British**	**Underrepresented**
1	b	b	b	b
2	.68 (2.67)	2.13 (2.61)	3.11 (2.05)	-.62 (3.79)
3	1.63 (2.89)	4.36 (3.11)	4.34 (2.40)	-.19 (3.87)
4	6.91* (2.68)	-3.07 (2.71)	-.73 (2.23)	6.05 (3.50)
5	-.65 (3.51)	.38 (2.60)	1.65 (2.14)	.13 (3.60)
6	3.18 (4.76)	-4.48 (5.40)	3.61 (3.47)	-7.62 (5.70)
7	3.24 (3.40)	-4.02 (3.02)	.50 (2.47)	-2.06 (4.14)
8	1.94 (3.70)	1.96 (5.47)	2.32 (3.57)	1.10 (5.83)

9	-.89 (3.00)	-.96 (2.78)	.22 (2.17)	1.09 (3.73)
10	3.06 (4.17)	-6.83 (3.99)	-2.97 (3.22)	1.59 (4.91)
11	[c]	-7.68 (5.62)	-4.18 (4.95)	.46 (6.39)
12	9.91* (3.50)	-.89 (2.99)	4.90 (2.54)	-.52 (4.19)
Baseline OLBI	.87[a]* (.30)	.18[a] (.28)	.40[a] (.23)	.84[a]* (.36)

Note. SE = Standard Error; Underrepresented Racial-Ethnic Group = Asian, Black, Chinese, multiracial, and Other; Junior Doctor = Foundation, Core Training, Higher Training, Vocational Training Scheme; **p* = < .05; ***p* = <.001.

[a] Dependent Variable: Post-Treatment OLBI.

[b] Cluster 1 was indicated as the reference group.

[c] Cluster 11 was excluded from one analysis as it became redundant.

Discussion

Summary of Main Results

This study investigated the use of machine learning to identify and validate subtypes of burnout and examine their predictive value for treatment responses to the MMSFL programme among NHS doctors and nurses. Twelve subtypes of burnout were identified using SOM, the linear model was selected based on its interpretability. The first hypothesis was supported as findings indicate that HCP burnout can appear in different subtypes and might present in more varied ways than established definitions (e.g., OLBI domains). Average baseline burnout levels varied across individual clusters (i.e., subtypes). However, a comparable range of mild to moderate burnout was observed across the training and test samples, providing preliminary evidence of generalisability of subtypes across nursing and medical professions. Clusters 3 and 10 in the training sample emerged as the most prevalent, showing the highest and lowest burnout severity, respectively. Initial regression analyses indicated that members of clusters 8, 10, and 12 (training sample) and cluster 10 (test sample) were associated with more favourable outcomes, whereas clusters 2 and 3 (test sample) showed less favourable outcomes relative to the reference group.

In further analyses, regression models were adjusted for baseline severity and the predictive power of clusters diminished in both samples. The second hypothesis was not supported as no clusters remained significantly associated with outcomes, suggesting subtypes independent of baseline severity did not influence differential treatment responses across the samples. Baseline OLBI severity remained the only significant predictor for the test sample; however, this pattern was not found in the training sample. Across samples, scatterplots demonstrated a positive relationship between baseline and post-treatment OLBI scores. This suggests that HCPs with more severe baseline burnout tended to report more severe levels after the MMSFL programme. Sensitivity analyses were run in the test sample, revealing differential treatment responses across clusters based on demographic and occupational factors. Junior doctors and doctors from underrepresented racial-ethnic identities reported more severe levels of burnout at baseline and post-treatment, even after adjusting for baseline severity. These factors were not tested in the training sample as the required data was not provided in the CPM 1 dataset.

Based on these findings, it is currently unclear how to personalise treatment for the MMSFL programme, at least based on the variables available in the samples. The twelve subtypes provide novel manifestations of burnout; however, their clinical utility in predicting treatment outcomes is limited once baseline severity is considered. Baseline severity did not consistently predict outcomes across samples. An explanation for this might be that subtypes and initial burnout severity levels were explaining overlapping variance in treatment outcomes, or the initial significant results for clusters were not robust (Kim, 2019). Baseline burnout differences may have also confounded the perceived advantages of the subtypes; it is possible that this is a function of where participants started and there being greater distance to travel on the OLBI for measurable improvement to be observed in those reporting more severe initial burnout. Demographic and occupational factors suggest differential treatment responses in certain subgroups. However, these findings should be interpreted with caution due to the lack of predictive value in clusters and baseline severity across samples. This highlights the importance of external cross-validation procedures in determining which clusters and other factors could be clinically meaningful.

Positioning Within the Current Evidence Base

In line with previous research, this study demonstrates the ongoing difficulty of reducing burnout among healthcare workforces. This fits with the wider literature as systematic reviews tend to report short-term benefits for individual-level interventions, highlighting limitations regarding long-term maintenance (Clough et al., 2017; Cohen et al., 2023; Tamminga et al., 2023). Evidence concerning DHIs remains inconsistent, with studies reporting minimal or no improvement in burnout outcomes (Barrett & Stewart, 2021; Ilola at al., 2024). Personal and cultural differences have also emerged as key barriers to the effectiveness of DHIs (Aye et al., 2024). A mixed methods study of HCPs (mainly doctors and nurses) in the UK and China identified mental health stigma, language barriers, and cultural preferences (e.g., favouring personal support networks) as obstacles to engagement with DHIs among Asian and Black groups (Zhang et al., 2025). These findings emphasise the need to consider the intersections of social identities (e.g., gender, racial-ethnic identity, cultural influences) when investigating the efficacy of interventions.

Other research indicates that current burnout measures may not sufficiently capture the complexity of burnout, making it difficult to measure subtle meaningful improvements in more severe presentations, which could underestimate therapeutic progress (Nadon et al., 2022; Leiter & Maslach, 2016). Global systematic and meta-analytic research among health professionals show that individuals presenting with more severe baseline burnout often exhibit greater comorbidity with depression and anxiety (Koutsimani et al., 2019; Ryan et al., 2023; Salvagioni et al., 2017). Moreover, increasing numbers of studies are investigating the potential overlap of emotional exhaustion as a feature of burnout and depression (Baptista et al., 2022). Taken together, these findings offer other possible explanations for the pattern of results in this study, emphasising how burnout affects individuals in a multitude of ways, specifically how severity of baseline burnout might limit observable improvement. This further highlights the importance of comprehensive screening measures to detect comorbidities and guide treatment planning

In addition, meta-analytic findings suggest that core domains of burnout (i.e., emotional exhaustion, depersonalisation, personal accomplishment) may respond differentially to well-established interventions. Reductions in emotional exhaustion are the strongest and most consistent findings documented in the literature for individual and organisational interventions (Maricuţoiu et al., 2014; Panagioti et al., 2017; West et al., 2016). Similarly, moderate short-term benefits are reported among HCPs in a recent Cochrane review following cognitive behavioural (e.g., coping skills, assertiveness, communication) and mindfulness (e.g., stress reduction, relaxation, goal-based skills) techniques (Tamminga et al., 2023). However, the evidence base for other domains is more limited and less reliable. This study demonstrates how burnout remains a multifaceted construct, impacting HCPs in various ways despite similar burnout features. The identified burnout subtypes did not predict differential treatment responses, indicating that the characteristics of the twelve subtypes did not provide a broader understanding of the efficacy of the MMSFL programme than reported in the initial trial using the core OLBI domains (Laker et al., 2023).

Strengths, Limitations, and Areas for Future Research

A main strength of this study was the use of SOM, a powerful exploratory and data driven method that reveals underlying patterns and natural relationships in complex multidimensional data (Kohonen, 2013). In contrast to hypothesis-driven statistical tests, SOM does not require predefined labels or assumptions, meaning it can identify subgroups (clusters of similar characteristics) that might remain undetected in standard methods (Galkin et al., 2022). The external-cross validation procedure was another methodological strength, which tested the generalisability of subtypes in independent datasets, providing a more rigorous assessment of the real-world applicability of the subtypes across clinical professions (Delgadillo, 2021). In addition, the sensitivity analyses assessed the robustness of findings, specifically what other factors might influence treatment responses. While a default reference group (cluster 1) in the regression models provided a useful baseline for comparison, it is important to acknowledge how other clusters may have yielded significant results if they were selected as the reference group. Future research could consider a more meaningful reference category (e.g., the most common cluster), ensuring it aligns with the research question, to enhance the precision and interpretability of outcomes.

In regard to the external cross-validation procedure, using different professions for the training and test samples was advantageous in reducing the risk of overfitting the burnout subtypes to profession-specific contexts, increasing the robustness and confidence in the underlying burnout patterns. However, well-established burnout measures may not equally capture the experiences in each group (e.g., role-specific workload pressures) which could have impacted how well the subtypes replicated across the professions and underestimated treatment response patterns in regression analyses. This offers another possible explanation of the inconsistent findings for treatment response across samples. While an alternative approach to partitioning the datasets as a 50:50 within-profession split could have mitigated some of these issues and identified more accurate subtype classifications for each profession, it would have restricted the ability to run sensitivity analyses within the doctor sample as smaller subgroupings of underrepresented racial-ethnic groups may not have been large enough to conduct an adequately powered analysis. In future research, larger and more diverse samples of different professions could be merged with half of the data randomly designated as the training and test datasets. While this may increase the stability of clusters across professions, it is important to consider the overall aim of the research and how this could impact generalisability compared to the other approaches.

Several other notable limitations were identified, including how burnout was measured. Established burnout severity levels were used (Leclercq et al., 2021); however, using a reliable change index could have examined whether there was a statistically reliable change in OLBI scores, increasing interpretability beyond severity levels. Despite this, the absence of a defined clinical cut-off helped to capture individual heterogeneity and enhanced the inclusivity of findings (Huynh et al., 2021). This was particularly important as burnout manifested in different ways across the subtypes. The self-report nature of the OLBI may have introduced response bias. For example, the underreporting of symptoms may have been an issue among doctors and nurses where resilience and competence is valued within the professions. While the data driven approach remains a strength of SOM, it can be difficult to interpret the reasons why individuals have been grouped together. Moreover, item-level data was used to visualise the clusters via bar charts to increase the meaningfulness of subtypes. As the burnout subtypes alone did not provide additional predictive value for treatment responses, future studies could examine a more extensive range of sociocultural and occupational factors (e.g., structural inequities, job grading, speciality) to better understand how these might predict outcomes to the MMSFL programme among diverse groups of HCPs.

Additionally, the regression models were adjusted for baseline severity. This step was critical in revealing the influence of initial burnout severity on treatment responses and can help to control the impact of regression to the mean, a statistical phenomenon where extreme scores (e.g., more severe initial burnout) are closer to the mean when measurements are repeated and might be mistaken for true treatment effects (Barnett et al., 2005). This enabled a more accurate and reliable understanding of how the MMSFL programme works across various levels of initial severity, increasing statistical power of the study. Conversely, the training sample did not provide data regarding racial-ethnic identity or job grading. This restricted the completion of sensitivity analyses in the CPM 1 dataset (i.e., sample of nurses), meaning the study was unable to test for differential treatment effects based on these factors across samples. Future research should be consistent in gathering demographic and occupational factors to increase the understanding of treatment responses within diverse healthcare workforces, which may also include allied HCPs and non-NHS settings (e.g., private, voluntary, and charitable sectors).

Future research would benefit from addressing limitations and broadening study findings. Incorporating multiple baseline measures would enable a broader understanding of how comorbid mental health difficulties such as depression and anxiety might influence subtype classification and treatment responses. Studies should also prioritise recruitment and analysis of more than one sample comprising HCPs with diverse racial-ethnic identities and job gradings across professions. This would help to better understand cultural and contextual factors that may shape burnout experiences and treatment effects to support the tailoring of NHS staff interventions. Longitudinal designs could also test whether extending the MMSFL programme would demonstrate meaningful improvement for cases with more severe initial burnout severity, using the OLBI as it has been validated in longitudinal studies with stability over multiple observations (Demerouti et al., 2003). Qualitative studies could explore the cultural relevance of the MMSFL programme, providing rich information about the lived experiences of a range of racial-ethnic identities to support the delivery of more equitable care (Bansal et al., 2022).

Clinical Implications

At this stage, findings in the present study suggest there would be no additional benefits to tailoring the MMSFL intervention (Laker et al., 2023) based on the 12 identified burnout subtypes. While preliminary evidence is shown for the generalisability of the subtypes based on similar mild to moderate baseline burnout levels across professions, it would be premature to speculate about how this might contribute to the conceptualisation of burnout until the clinical utility of subtypes regarding treatment response is confirmed in future research. This may be supported by considering other approaches to partitioning samples (e.g., within-profession split) to undertake external cross-validation. Baseline severity did not consistently predict treatment outcomes which differs from other psychotherapy literature (Gold et al., 2024). However, descriptive analyses suggest more severe initial burnout was related to less observable improvement after the MMSFL programme. Baseline severity should therefore be held in mind when designing the intensity and duration of burnout interventions in NHS settings. Routine use of validated burnout measures such as the OLBI would help to identify HCPs experiencing more severe initial burnout, indicating that an extended intervention or follow-up period, and possible 1:1 therapy, might be required. Using mental health screening measures such as the Patient Health Questionnaire-9 for depressive symptoms could also support decisions regarding interventions for comorbid presentations.

Additionally, there is preliminary evidence indicating more baseline and post-treatment severity in some subgroups of junior doctors and doctors from underrepresented racial-ethnic identities, highlighting the need for more targeted and equitable approaches. Recent NHS workforce statistics show 25.7% of staff identify from underrepresented racial-ethnic groups (NHS Digital, 2022), indicating the sample of underrepresented doctors in this study was representative of the target population. These groups face unique systemic stressors such as racial discrimination, intense workloads, lower autonomy, and language barriers (Lawrence et al., 2021; Zhang et al., 2025). Rather than a universal approach to the MMSFL programme, NHS services adopting this intervention could explore options for stratified care that considers baseline severity, as well as culturally sensitive care pathways. This may help to meet preferences based on racial-ethnic identity, religion, or spirituality, including culturally sensitive resources and psychoeducation. Integrating these factors has the potential to increase the acceptability of burnout interventions in the increasingly diverse and growing NHS workforce.

Finally, the MMSFL programme could be offered alongside organisational approaches such as structured supervision, workload management, and mentoring programmes (Walsh et al., 2019). While organisational interventions have tended to show mixed effects due to focusing on one domain of burnout, they are able to address systemic factors in environments where HCPs have less control over the stressors they encounter (Mostafa et al., 2025). Building on the previous findings in Laker et al. (2023), the MMSFL programme appears to be an intervention that is more generalisable across a range of mild to moderate burnout profiles as there were no significant differences between subtypes when baseline severity was controlled for in the primary analyses. That said, NHS services ought to also consider how to tailor workplace environments to demographic and occupational needs alongside the delivery of the MMSFL programme. For example, racially diverse and supportive environments have been shown to be protective against burnout (Bafna et al., 2025), which is particularly important for groups that may face more disparities in the workplace, such as underrepresented racial-ethnic identities, or junior doctors and newly-qualified staff who may be at increased risk of experiencing early career burnout.

Conclusion

To conclude, we do not currently have a robust understanding about how to personalise treatment for the MMSFL programme, at least based on the samples examined in this study. Preliminary findings offer valuable insights regarding baseline burnout severity, which appears to be associated with outcomes among subgroups of HCPs based on training grade (i.e., junior doctors) and racial-ethnic identity (i.e., doctors who self-identified from underrepresented groups), rather than burnout subtype alone. This warrants further investigation in future studies to confirm if the findings replicate in other diverse samples, which would increase the robustness and clinical utility of the 12 identified subtypes and treatment effects for particular subgroups based on demographic and occupational factors. Moreover, this could support tailoring of the MMSFL programme to job roles or personal preferences, providing future research opportunities to examine treatment responses across novel burnout profiles and more diverse subgroups of HCPs with careful consideration to the impact of baseline severity.

References

Aye, L. M., Tan, M. M., Schaefer, A., Thurairajasingam, S., Geldsetzer, P., Soon, L. K., Reininghaus, U., Bärnighausen, T., & Su, T. T. (2024). Self-help digital mental health intervention in improving burnout and mental health outcomes among healthcare workers: A narrative review. *Digital Health*, *10*, 1-19. https://doi.org/10.1177/20552076241278313

Azari, B., Westlin, C., Satpute, A. B., Hutchinson, J. B., Kragel, P. A., Hoemann, K., Khan, Z., Wormwood, J. B., Quigley, K. S., Erdogmus, D., Dy, J. G., Brooks, D. H., & Barrett, L. F. (2020). Comparing supervised and unsupervised approaches to emotion categorization in the human brain, body, and subjective experience. *Scientific Reports, 10*(1). https://doi.org/10.1038/s41598-020-77117-8

Bafna, T., Malik, M., Choudhury, M. C., Hu, W., Weston, C. M., Weeks, K. R., Connors, C., Burhanullah, M. H., Everly, G., Michtalik, H. J., & Wu, A. W. (2025). Systematic review of well-being interventions for minority healthcare workers. *Frontiers in Medicine, 12*. Article e1531090. https://doi.org/10.3389/fmed.2025.1531090

Bansal, N., Karlsen, S., Sashidharan, S. P., Cohen, R., Chew-Graham, C. A., & Malpass, A. (2022). Understanding ethnic inequalities in mental healthcare in the UK: A meta-ethnography. *PLoS Medicine, 19*(12), Article e1004139. https://doi.org/10.1371/journal.pmed.1004139

Baptista, M. N., Hauck-Filho, N., & Cardoso, H. F. (2022). The overlap between burnout and depression through a different lens: A multi-method study. *Journal of Affective Disorders Reports, 10,* Article 100437. https://doi.org/10.1016/j.jadr.2022.100437

Barnett, A. G., Van Der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology, 34*(1), 215-220. https://doi.org/10.1093/ije/dyh299

Barrett, K., & Stewart, I. (2021). A preliminary comparison of the efficacy of online acceptance and Commitment Therapy (ACT) and Cognitive Behavioural Therapy (CBT) stress management interventions for social and healthcare workers. *Health and Social Care in the Community*, *29*(1), 113–126. https://doi.org/10.1111/hsc.13074

Beacham, A. O., King, A. W., & Nash, B. F. (2023). Psychologists' role in addressing healthcare provider burnout and well-being. *Journal of Health Service Psychology*, *49*(2), 63–75. https://doi.org/10.1007/s42843-023-00083-y

Bowen, F., & Siegler, J. (2024). Self-organizing maps: a novel approach to identify and map business clusters. *Journal of Management Analytics, 11*(2), 228-246. https://doi.org/10.1080/23270012.2024.2306628

Breiman, L. (2001). Statistical modelling: The two cultures (with comments and a rejoinder by the author). *Statistical Science 16*(3), 199-231. https://doi.org/doi:10.1214/ss/1009213726

Clough, B. A., March, S., Chan, R. J., Casey, L. M., Phillips, R., & Ireland, M. J. (2017). Psychosocial interventions for managing occupational stress and burnout among medical doctors: a systematic review. *Systematic Reviews*, *6*(1). https://doi.org/10.1186/s13643-017-0526-3

Cohen, C., Pignata, S., Bezak, E., Tie, M., & Childs, J. (2023). Workplace interventions to improve well-being and reduce burnout for nurses, physicians and allied healthcare professionals: a systematic review. *BMJ Open, 13*(6), e071203. https://doi.org/10.1136/bmjopen-2022-071203

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, Z. D., DeRubeis, R. J., Hayes, R., Watkins, E. R., Lewis, G., Byng, R., Byford, S., Crane, C., Kuyken, W., Dalgleish, T., & Schweizer, S. (2022). The development and internal evaluation of a predictive model to identify for whom mindfulness-based cognitive therapy offers superior relapse prevention for recurrent depression versus maintenance antidepressant medication. *Clinical Psychological Science, 11*(1), 59-76. https://doi.org/10.1177/21677026221076832

Collins, G. S., Moons, K. G. M., Dhiman, P., Riley, R. D., Beam, A. L., Van Calster, B., Ghassemi, M., Liu, X., Reitsma, J. B., Van Smeden, M., Boulesteix, A., Camaradou, J. C., Celi, L. A., Denaxas, S., Denniston, A. K., Glocker, B., Golub, R. M., Harvey, H., Heinze, G., . . . Logullo, P. (2024). TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ, 385,* Article e078378. https://doi.org/10.1136/bmj-2023-078378

CPM Trial Research Group. (2023). *Clinical trial of an intervention to support medical doctors with occupational burnout*. https://doi.org/10.1186/ISRCTN14947225

De Hert, S. (2020). Burnout in healthcare workers: Prevalence, impact and preventative strategies. *Local and Regional Anesthesia, 13,* 171-183. https://doi.org/10.2147/LRA.S240564

Delgadillo, J. (2021). Machine learning: A primer for psychotherapy researchers. *Psychotherapy Research*, *31*(1), 1-4. https://doi.org/10.1080/10503307.2020.1859638

Delgadillo, J., & Atzil-Slonim. (2022). Artificial intelligence, machine learning and mental health. In *Elsevier eBooks* (pp. 132–142). https://doi.org/10.1016/b978-0-323-91497-0.00177-6

Delgadillo, J., Saxon, D., & Barkham, M. (2018). Associations between therapists’ occupational burnout and their patients’ depression and anxiety treatment outcomes. *Depression and Anxiety, 35*(9), 844-850. https://doi.org/10.1002/da.22766

Demerouti, E., Bakker, A. B., Nachreiner, F., & Schaufeli, W. B. (2001). The job demands-resources model of burnout. *Journal of Applied Psychology, 86*(3), 499-512. https://doi.org/10.1037/0021-9010.86.3.499

Demerouti, E., Bakker, A. B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments: A multitrait-multimethod analysis. *European Journal of Psychological Assessment, 19*(1), 12–23. https://doi.org/10.1027/1015-5759.19.1.12

Dincer, B., & Inangil, D. (2021). The effect of emotional freedom techniques on nurses' stress, anxiety, and burnout levels during the covid-19 pandemic: A randomized controlled trial. *EXPLORE, 17*(2), 109–114. https://doi.org/10.1016/j.explore.2020.11.012

Fava, M., Rush, A. J., Alpert, J. E., Carmin, C. N., Balasubramani, G., Wisniewski, S. R., Trivedi, M. H., Biggs, M. M., & Shores-Wilson, K. (2006). What clinical and symptom features and comorbid disorders characterize outpatients with anxious major depressive disorder: a replication and extension. *The Canadian Journal of Psychiatry*, *51*(13), 823–835. https://doi.org/10.1177/070674370605101304

Fava, M., Uebelacker, L. A., Alpert, J. E., Nierenberg, A. A., Pava, J. A., & Rosenbaum, J. F. (1997). Major depressive subtypes and treatment response. *Biological Psychiatry, 42*(7), 568-576. https://doi.org/10.1016/S0006-3223(96)00440-4

Galkin, F., Kochetov, K., Keller, M., Zhavoronkov, A., & Etcoff, N. (2022). Optimizing future well-being with artificial intelligence: self-organizing maps (SOMs) for the identification of islands of emotional stability. *Aging, 14*(12), 4935-4958. https://doi.org/10.18632/aging.204061

General Medical Council. (2023). *National training survey 2023 results.* [GMC/NTSR2023/0723]. https://www.gmc-uk.org/-/media/documents/national-training-survey-2023-initial-findings-report_pdf-101939815.pdf

Ghahramani, S., Lankarani, K. B., Yousefi, M., Heydari, K., Shahabi, S., & Azmand, S. (2021). A systematic review and meta-analysis of burnout among healthcare workers during COVID-19. *Frontiers in Psychiatry, 12,* Article e758849. https://doi.org/10.3389/fpsyt.2021.758849

Gold, A. K., Rabideau, D. J., Nolte, D., Faria, C. M., Deng, S. Y., George, N., Boccagno, C., Temes, C. M., Kamali, M., Akpolat, N., Nierenberg, A. A., & Sylvia, L. G. (2024). Does baseline psychiatric symptom severity predict well-being improvement in low-intensity mindfulness interventions? *Psychiatry Research Communications, 4*(3), Article 100182. https://doi.org/10.1016/j.psycom.2024.100182

Halbesleben, J. R. B., & Demerouti, E. (2005). The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. *Work & Stress, 19*(3), 208–220. https://doi.org/10.1080/02678370500340728

Hall, L. H., Johnson, J., Heyhoe, J., Watt, I., Anderson, K., & O'Connor, D. B. (2020). Exploring the impact of primary care physician burnout and well-being on patient care: A focus group study. *Journal of Patient Safety, 16*(4), e278-e283. https://doi.org/10.1097/PTS.0000000000000438

Harvey, S. B., Epstein, R. M., Glozier, N., Petrie, K., Strudwick, J., Gayed, A., Dean, K., & Henderson, M. (2021). Mental illness and suicide among physicians. *The Lancet*, *398*(10303), 920–930. https://doi.org/10.1016/s0140-6736(21)01596-8

Health Education England. (2022). *The F3 phenomenon: Exploring a new norm and its implications.* https://www.hee.nhs.uk/sites/default/files/documents/F3_Phenomenon_Final.pdf

Holmes, J. (2022, October 1). *The NHS nursing workforce – have the floodgates opened?* The Kind's Fund. https://www.kingsfund.org.uk/insight-and-analysis/blogs/nhs-nursing-workforce?

Huibers, M. J. H., Cohen, Z. D., Lemmens, L. H. J. M., Arntz, A., Peeters, F. P. M. L., Cuijpers, P., & DeRubeis, R. J. (2016). Correction: Predicting optimal outcomes in cognitive therapy or interpersonal psychotherapy for depressed individuals using the Personalized Advantage Index approach. *PLoS ONE, 11*(2), Article e0140771. https://doi.org/10.1371/journal.pone.0140771

Huynh, A. N., Besse, C., Mediouni, Z., May, E. E., Shoman, Y., Hansez, I., & Canu, I. G. (2021). Diagnostic performances of an occupational burnout detections method designed for healthcare professionals. *International Journal of Environmental Research and Public Health, 18*(23), Article 12300. https://doi.org/10.3390/ijerph182312300

Ilola, T., Malmisalo, M., Laukka, E., Lehtiniemi, H., Pölkki, T., Kääriäinen, M., He, H., & Kanste, O. (2024). The effectiveness of digital solutions in improving nurses' and healthcare professionals' mental well-being: a systematic review and meta-analysis. *Journal of Research in Nursing*, *29*(2), 97–109. https://doi.org/10.1177/17449871241226914

Kanzler, K. E., Goodie, J. L., Hunter, C. L., Glotfelter, M. A., & Bodart, J. J. (2013). From colleague to patient: Ethical challenges in integrated primary care. *Families, Systems, & Health, 31*(1), 41-48. https://doi.org/10.1037/a0031853

Kappelmann, N., Rein, M., Fietz, J., Mayberg, H. S., Craighead, W. E., Dunlop, B. W., Nemeroff, C. B., Keller, M., Klein, D. N., Arnow, B. A., Husain, N., Jarrett, R. B., Vittengl, J. R., Menchetti, M., Parker, G., Barber, J. P., Bastos, A. G., Dekker, J., Peen, J., . . . Kopf-Beck, J. (2020). Psychotherapy or medication for depression? Using individual symptom meta-analyses to derive a Symptom-Oriented Therapy (SOrT) metric for a personalised psychiatry. *BMC Medicine*, *18*(1). https://doi.org/10.1186/s12916-020-01623-9

Kiang, M. Y. (2001). Extending the Kohonen self-organizing map networks for clustering analysis. *Computational Statistics & Data Analysis, 38*(2), 161-180. https://doi.org/10.1016/S0167-9473(01)00040-8

Kiang, M. Y., Hu, M. Y., & Fisher, D. M. (2007). The effect of sample size on the extended self-organizing map network—A market segmentation application. *Computational statistics & data analysis, 51*(12), 5940-5948. https://doi.org/10.1016/j.csda.2006.11.011

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology, 72*(6), 558-569. https://doi.org/10.4097/kja.19087

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks, 37*, 52-65. https://doi.org/10.1016/j.neunet.2012.09.018

Koutsimani, P., Montgomery, A., & Georganta, K. (2019). The relationship between burnout, depression, and anxiety: A systematic review and meta-analysis. *Frontiers in Psychology, 10,* Article 284. https://doi.org/10.3389/fpsyg.2019.00284

Laker, V., Simmonds-Buckley, M., Delgadillo, J., Palmer, L., & Barkham, M. (2023). Pragmatic randomized controlled trial of the Mind Management Skills for Life Programme as an intervention for occupational burnout in mental healthcare professionals. *Journal of Mental Health, 32*(4), 752-760. https://doi.org/10.1080/09638237.2023.2182423

Lamers, F., De Jonge, P., Nolen, W. A., Smit, J. H., Zitman, F. G., Beekman, A. T. F., & Penninx, B. W. J. H. (2010). Identifying depressive subtypes in a large cohort study: results from the Netherlands Study of Depression and Anxiety (NESDA). *The Journal of Clinical Psychiatry*, *71*(12), 1582–1589. https://doi.org/10.4088/jcp.09m05398blu

Lawrence, J. A., Davis, B. A., Corbette, T., Hill, E. V., Williams, D. R., & Reede, J. Y. (2021). Racial/ethnic differences in burnout: a systematic review. *Journal of Racial and Ethnic Health Disparities, 9*(1), 257-269. https://doi.org/10.1007/s40615-020-00950-0

Leclercq, C., Braeckman, L., Firket, P., Babic, A., & Hansez, I. (2021). Interest of a joint use of two diagnostic tools of burnout: comparison between the oldenburg burnout inventory and the early detection tool of burnout completed by physicians. *International Journal of Environmental Research and Public Health*, *18*(19), Article 10544. https://doi.org/10.3390/ijerph181910544

Lee, H., Kuo, C., Chien, T., & Wang, Y. (2016). A meta-analysis of the effects of coping strategies on reducing nurse burnout. *Applied Nursing Research, 31*, 100–110. https://doi.org/10.1016/j.apnr.2016.01.001

Lee, J., Lim, N., Yang, E., & Lee, S. M. (2011). Antecedents and consequences of three dimensions of burnout in psychotherapists: A meta-analysis. *Professional Psychology: Research and Practice, 42*(3), 252–258. https://doi.org/10.1037/a0023319

Leiter, M. P., & Maslach, C. (2016). Latent burnout profiles: A new approach to understanding the burnout experience. *Burnout Research, 3*(4), 89-100. https://doi.org/10.1016/j.burn.2016.09.001

Li, L. Z., Yang, P., Singer, S. J., Pfeffer, J., Mathur, M. B., & Shanafelt, T. (2024). Nurse burnout and patient safety, satisfaction, and quality of care. *JAMA Network Open*, *7*(11), Article e2443059. https://doi.org/10.1001/jamanetworkopen.2024.43059

Lutz, W., de Jong, K., Rubel, J. A., & Delgadillo, J. (2021). Measuring, predicting, and tracking change in psychotherapy. *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed.). Wiley.

Mamun, A., & Paul, S. (2023). Model selection in generalized linear models. *Symmetry*, *15*(10), Article 1905. https://doi.org/10.3390/sym15101905

Maricuțoiu, L. P., Sava, F. A., & Butta, O. (2014). The effectiveness of controlled interventions on employees' burnout: A meta-analysis. *Journal of Occupational and Organizational Psychology, 89*(1), 1-27. https://doi.org/10.1111/joop.12099

Maslach, C., & Leiter, M. P. (2016). Understanding the burnout experience: recent research and its implications for psychiatry. *World Psychiatry, 15*(2), 103-111. https://doi.org/10.1002/wps.20311

Medisauskaite, A., Silkens, M., & Rich, A. (2023). A national longitudinal cohort study of factors contributing to UK medical students' mental ill-health symptoms. *General Psychiatry, 36*(2). https://doi.org/10.1136/gpsych-2022-101004

Moe-Byrne, T., Shepherd, J., Merecz-Kot, D., Sinokki, M., Naumanen, P., Roijen, L. H., & Van Der Feltz-Cornelis, C. (2022). Effectiveness of tailored digital health interventions for mental health at the workplace: A systematic review of randomised controlled trials. *PLOS Digital Health*, *1*(10), Article e0000123. https://doi.org/10.1371/journal.pdig.0000123

Mostafa, I., Onwumere, J., & Wood, L. (2025). A systematic review and thematic synthesis of healthcare professionals' experiences of racism in the workplace and the support they receive. *Journal of Workplace Behavioral Health,* 1-29. https://doi.org/10.1080/15555240.2025.2519298

Musil, R., Seemüller, F., Meyer, S., Spellmann, I., Adli, M., Bauer, M., Kronmüller, K., Brieger, P., Laux, G., Bender, W., Heuser, I., Fisher, R., Gaebel, W., Schennach, R., Möller, H., & Riedel, M. (2018). Subtypes of depression and their overlap in a naturalistic inpatient sample of major depressive disorder. *International Journal of Methods in Psychiatric Research*, *27*(1) Article e1569. https://doi.org/10.1002/mpr.1569

Nadon, L., De Beer, L. T., & Morin, A. J. S. (2022). Should burnout be conceptualized as a mental disorder? *Behavioral Sciences, 12*(3), Article 82. https://doi.org/10.3390/bs12030082

Nash, S (2025, July 31). *From burnout to belief: reflections on reforming the NHS from within.* The King's Fund. https://www.kingsfund.org.uk/insight-and-analysis/long-reads/burnout-belief-reforming-the-nhs

National Health Service Digital. (2022, September 29). *NHS Workforce Statistics - June 2022 (including selected provisional statistics for July 2022).* https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/june-2022

National Health Service England. (2023, June). *NHS Long Term Workforce Plan.* https://www.england.nhs.uk/long-read/accessible-nhs-long-term-workforce-plan/

National Institute for Health and Care Excellence. (2022). *Mental Wellbeing at Work* [NICE Guidelines No. 212]. https://www.nice.org.uk/guidance/ng212

Núñez, C., Delgadillo, J., Barkham, M., & Behn, A. (2024). Understanding symptom profiles of depression with the PHQ-9 in a community sample using network analysis. *European Psychiatry*, *67*(1), Article e50. https://doi.org/10.1192/j.eurpsy.2024.1756

O'Connor, K., Neff, M. D., & Pitman, S. (2018). Burnout in mental health professionals: A systematic review and meta-analysis of prevalence and determinants. *European Psychiatry, 53,* 74-99. https://doi.org/10.1016/j.eurpsy.2018.06.003

Panagioti, M., Panagopoulou, E., Bower, P., Lewith, G., Kontopantelis, E., Chew-Graham, C., Dawson, S., Van Marwijk, H., Geraghty, K., & Esmail, A. (2017). Controlled interventions to reduce burnout in physicians. *JAMA Internal Medicine*, *177*(2), 195-205. https://doi.org/10.1001/jamainternmed.2016.7674

Paterson, C., Leduc, C., Maxwell, M., Aust, B., Strachan, H., O'Connor, A., Tsantila, F., Cresswell-Smith, J., Purebl, G., Winter, L., Fanaj, N., Doukani, A., Hogg, B., Corcoran, P., D'Alessandro, L., Mathieu, S., Hegerl, U., Arensman, E., Greiner, B. A., . . . Orchard, W. (2024). Barriers and facilitators to implementing workplace interventions to promote mental health: qualitative evidence synthesis. *Systematic Reviews, 13*, Article 152. https://doi.org/10.1186/s13643-024-02569-2

Peters, S. (2012). *The chimp paradox: The mind management programme to help you achieve success, confidence and happiness.* Vermilion.

Ramspek, C. L., Jager, K. J., Dekker, F. W., Zoccali, C. & Van Diepen, M. (2020). External validation of prognostic models: what, why, how, when and where? *Clinical Kidney Journal, 14*(1), 49-58. https://doi.org/10.1093/ckj/sfaa188

Ravalier, J. M., McVicar, A., & Boichat, C. (2020). Work stress in NHS employees: A mixed-methods study. *International Journal of Environmental Research and Public Health*, *17*(18), Article 6464. https://doi.org/10.3390/ijerph17186464

Ryan, E., Hore, K., Power, J., & Jackson, T. (2023). The relationship between physician burnout and depression, anxiety, suicidality and substance abuse: A mixed methods systematic review. *Frontiers in Public Health, 11,* Article e1133484. https://doi.org/10.3389/fpubh.2023.1133484

Salciccioli, J. D., Crutain, Y., Komorowski, M., & Marshall, D. C. (2016). Sensitivity analysis and model variation. In MIT Critical Data (Ed.), *Secondary analysis of electronic health records* (1st ed., pp. 263-271). Springer International Publishing. https://doi.org/10.1007/978-3-319-43742-2_17

Salvagioni, D. A. J., Melanda, F. N., Mesas, A. E., González, A. D., Gabani, F. L., & De Andrade, S. M. (2017). Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies. *PLoS ONE, 12*(10), 1–29. https://doi.org/10.1371/journal.pone.0185781

Scala, J. J., Ganz, A. B., Snyder, M. P. (2022). Precision medicine approaches to mental health care. *Physiology, 38,* 82-98. https://doi.org/doi:10.1152/physiol.00013.2022

Simmonds-Buckley, M., Catarino, A., & Delgadillo, J. (2021). Depression subtypes and their response to cognitive behavioral therapy: A latent transition analysis. *Depression and Anxiety*, *38*(9), 907–916. https://doi.org/10.1002/da.23161

Tamminga, S. J., Emal, L. M., Boschman, J. S., Levasseur, A., Thota, A., Ruotsalainen, J. H., Schelvis, R. M., Nieuwenhuijsen, K., & Van Der Molen, H. F. (2023). Individual-level interventions for reducing occupational stress in healthcare workers. *Cochrane Database of Systematic Reviews, Library*, *2023*(5). https://doi.org/10.1002/14651858.cd002892.pub6

Taris, T. W. (2006). Is there a relationship between burnout and objective performance? A critical review of 16 studies. *Work and Stress, 20*(4), 316–334. https://doi.org/10.1080/02678370601065893

Van Dam, A. (2021). A clinical perspective on burnout: diagnosis, classification, and treatment of clinical burnout. *European Journal of Work and Organizational Psychology, 30*, 732-741. https://doi.org/10.1080/1359432X.2021.1948400

Walsh, A. L., Lehmann, S., Zabinski, J., Truskey, M., Purvis, T., Gould, N. F., Stagno, S., & Chisolm, M. S. (2019). Interventions to prevent and reduce burnout among undergraduate and graduate medical education trainees: A systematic review. *Academic Psychiatry, 43*(4), 386–395. https://doi.org/10.1007/s40596-019-01023-z

West, C. P., Dyrbye, L. N., & Shanafelt, T. D. (2018). Physician burnout: contributors, consequences and solutions. *Journal of Internal Medicine, 283*(6), 516-529. https://doi.org/10.1111/joim.12752

West, C. P., Dyrbye, L. N., Erwin, P. J., & Shanafelt, T. D. (2016). Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. *The Lancet, 388*(10057), 2272-2281. https://doi.org/10.1016/S0140-6736(16)31279-X

Woo, T., Ho, R., Tang, A., & Tam, W. (2020). Global prevalence of burnout symptoms among nurses: A systematic review and meta-analysis. *Journal of Psychiatric Research, 123*, 9-20. https://doi.org/10.1016/j.jpsychires.2019.12.015

World Health Organization. (2019, May 28). *Burn-out an "occupational phenomenon": International Classification of Diseases.* https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases

Yin, H. (2008). The self-organizing maps: background, theories, extensions and applications. In J. Fulcher & L.C. Lain (Eds.), *Computational intelligence: A compendium* (pp.715-762). Springer.

Zhang, Z., Sun, S., Moradbakhti, L., Hall, A., Mougenot, C., Chen, J., & Calvo, R. A. (2025). Health care professionals' engagement with digital mental health interventions in the United Kingdom and China: Mixed methods study on engagement factors and design implications. *JMIR Mental Health*, *12*, Article e67190. https://doi.org/10.2196/67190

Appendices

Appendix A

Ethical Approval Letter

The University Of Sheffield.

Downloaded: 04/10/2024
Approved: 14/02/2024

Leanne Newsham
Registration number: 220237996
Psychology
Programme: Doctor of Clinical Psychology (DClinPsy) Programme

Dear Leanne

PROJECT TITLE: Identifying Occupational Burnout Subtypes in Healthcare Professionals Accessing the Mind Management Skills for Life Programme Intervention
APPLICATION: Reference Number 058389

This letter confirms that you have signed a University Research Ethics Committee-approved self-declaration to confirm that your research will involve only existing research, clinical or other data that has been robustly anonymised. You have judged it to be unlikely that this project would cause offence to those who originally provided the data, should they become aware of it.

As such, on behalf of the University Research Ethics Committee, I can confirm that your project can go ahead on the basis of this self-declaration.

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since full ethical review may be required.

Yours sincerely

Department Of Psychology Research Ethics Committee
Departmental Ethics Administrator

Appendix B

TRIPOD-AI Checklist (Collins et al., 2024)

TRIPOD+AI

Version: 11-January-2024

Section/Topic	Item	Development / evaluation[1]	Checklist item	Reported on page
TITLE				
Title	1	D;E	Identify the study as developing or evaluating the performance of a multivariable prediction model, the target population, and the outcome to be predicted	Page 67
ABSTRACT				
Abstract	2	D;E	See TRIPOD+AI for Abstracts checklist	Page 68
INTRODUCTION				
Background	3a	D;E	Explain the healthcare context (including whether diagnostic or prognostic) and rationale for developing or evaluating the prediction model, including references to existing models	Page 69-73
	3b	D;E	Describe the target population and the intended purpose of the prediction model in the context of the care pathway, including its intended users (e.g., healthcare professionals, patients, public)	Page 69-73
	3c	D;E	Describe any known health inequalities between sociodemographic groups	N/A
Objectives	4	D;E	Specify the study objectives, including whether the study describes the development or validation of a prediction model (or both)	Page 73-74
METHODS				
Data	5a	D;E	Describe the sources of data separately for the development and evaluation datasets (e.g., randomised trial, cohort, routine care or registry data), the rationale for using these data, and representativeness of the data	Page 74-75
	5b	D;E	Specify the dates of the collected participant data, including start and end of participant accrual; and, if applicable, end of follow-up	Page 74-75
Participants	6a	D;E	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including the number and location of centres	Page 74-76
	6b	D;E	Describe the eligibility criteria for study participants	Page 75-76
	6c	D;E	Give details of any treatments received, and how they were handled during model development or evaluation, if relevant	Page 76-77
Data preparation	7	D;E	Describe any data pre-processing and quality checking, including whether this was similar across relevant sociodemographic groups	Page 80
Outcome	8a	D;E	Clearly define the outcome that is being predicted and the time horizon, including how and when assessed, the rationale for choosing this outcome, and whether the method of outcome assessment is consistent across sociodemographic groups	Page 77
	8b	D;E	If outcome assessment requires subjective interpretation, describe the qualifications and demographic characteristics of the outcome assessors	N/A
	8c	D;E	Report any actions to blind assessment of the outcome to be predicted	N/A
Predictors	9a	D	Describe the choice of initial predictors (e.g., literature, previous models, all available predictors) and any pre-selection of predictors before model building	Page 79-83; Appendix C
	9b	D;E	Clearly define all predictors, including how and when they were measured (and any actions to blind assessment of predictors for the outcome and other predictors)	Page 80-83
	9c	D;E	If predictor measurement requires subjective interpretation, describe the qualifications and demographic characteristics of the predictor assessors	N/A
Sample size	10	D;E	Explain how the study size was arrived at (separately for development and evaluation), and justify that the study size was sufficient to answer the research question. Include details of any sample size calculation	Page 79-80
Missing data	11	D;E	Describe how missing data were handled. Provide reasons for omitting any data	Page 80
Analytical methods	12a	D	Describe how the data were used (e.g., for development and evaluation of model performance) in the analysis, including whether the data were partitioned, considering any sample size requirements	Page 80-83
	12b	D	Depending on the type of model, describe how predictors were handled in the analyses (functional form, rescaling, transformation, or any standardisation).	Page 80-83
	12c	D	Specify the type of model, rationale[2], all model-building steps, including any hyperparameter tuning, and method for internal validation	Page 80-83
	12d	D;E	Describe if and how any heterogeneity in estimates of model parameter values and model performance was handled and quantified across clusters (e.g., hospitals, countries). See TRIPOD-Cluster for additional considerations[3]	Page 80-83
	12e	D;E	Specify all measures and plots used (and their rationale) to evaluate model performance (e.g., discrimination, calibration, clinical utility) and, if relevant, to compare multiple models	Page 80-83
	12f	E	Describe any model updating (e.g., recalibration) arising from the model evaluation, either overall or for particular sociodemographic groups or settings	N/A
	12g	E	For model evaluation, describe how the model predictions were calculated (e.g., formula, code, object, application programming interface)	Page 80-83
Class imbalance	13	D;E	If class imbalance methods were used, state why and how this was done, and any subsequent methods to recalibrate the model or the model predictions	N/A
Fairness	14	D;E	Describe any approaches that were used to address model fairness and their rationale	N/A
Model output	15	D	Specify the output of the prediction model (e.g., probabilities, classification). Provide details and rationale for any classification and how the thresholds were identified	Page 80-83

TRIPOD+AI

Version: 11-January-2024

Training versus evaluation	16	D;E	Identify any differences between the development and evaluation data in healthcare setting, eligibility criteria, outcome, and predictors	N/A
Ethical approval	17	D;E	Name the institutional research board or ethics committee that approved the study and describe the participant-informed consent or the ethics committee waiver of informed consent	Page 74
OPEN SCIENCE				
Funding	18a	D;E	Give the source of funding and the role of the funders for the present study	Page 74
Conflicts of interest	18b	D;E	Declare any conflicts of interest and financial disclosures for all authors	Page ii
Protocol	18c	D;E	Indicate where the study protocol can be accessed or state that a protocol was not prepared	Page ii
Registration	18d	D;E	Provide registration information for the study, including register name and registration number, or state that the study was not registered	Page 75
Data sharing	18e	D;E	Provide details of the availability of the study data	Page ii
Code sharing	18f	D;E	Provide details of the availability of the analytical code[4]	Page ii
PATIENT & PUBLIC INVOLVEMENT				
Patient & Public Involvement	19	D;E	Provide details of any patient and public involvement during the design, conduct, reporting, interpretation, or dissemination of the study or state no involvement.	N/A
RESULTS				
Participants	20a	D;E	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Page 74-75
	20b	D;E	Report the characteristics overall and, where applicable, for each data source or setting, including the key dates, key predictors (including demographics), treatments received, sample size, number of outcome events, follow-up time, and amount of missing data. A table may be helpful. Report any differences across key demographic groups.	Page 77-80
	20c	E	For model evaluation, show a comparison with the development data of the distribution of important predictors (demographics, predictors, and outcome).	N/A
Model development	21	D;E	Specify the number of participants and outcome events in each analysis (e.g., for model development, hyperparameter tuning, model evaluation)	Page 80-83
Model specification	22	D	Provide details of the full prediction model (e.g., formula, code, object, application programming interface) to allow predictions in new individuals and to enable third-party evaluation and implementation, including any restrictions to access or re-use (e.g., freely available, proprietary)[5]	Page 83-90
Model performance	23a	D;E	Report model performance estimates with confidence intervals, including for any key subgroups (e.g., sociodemographic). Consider plots to aid presentation.	Page 83-95
	23b	D;E	If examined, report results of any heterogeneity in model performance across clusters. See TRIPOD Cluster for additional details[3].	Page 90-95
Model updating	24	E	Report the results from any model updating, including the updated model and subsequent performance	N/A
DISCUSSION				
Interpretation	25	D;E	Give an overall interpretation of the main results, including issues of fairness in the context of the objectives and previous studies	Page 95-97
Limitations	26	D;E	Discuss any limitations of the study (such as a non-representative sample, sample size, overfitting, missing data) and their effects on any biases, statistical uncertainty, and generalizability	Page 98-100
Usability of the model in the context of current care	27a	D	Describe how poor quality or unavailable input data (e.g., predictor values) should be assessed and handled when implementing the prediction model	N/A
	27b	D	Specify whether users will be required to interact in the handling of the input data or use of the model, and what level of expertise is required of users	N/A
	27c	D;E	Discuss any next steps for future research, with a specific view to applicability and generalizability of the model	Page 98-102

Appendix C

OLBI (Demerouti et al., 2001)

oldenburg burnout inventory

name: *date:*

Instructions: Below you find a series of statements with which you may agree or disagree. Using the scale, please indicate the degree of your agreement by selecting the number that corresponds with each statement.

		strongly agree	*agree*	*disagree*	*strongly disagree*
1.	I always find new and interesting aspects in my work *(D)*	*1*	*2*	*3*	*4*
2.	There are days when I feel tired before I arrive at work *(E.R.)*	*1*	*2*	*3*	*4*
3.	It happens more and more often that I talk about my work in a negative way *(D.R)*	*1*	*2*	*3*	*4*
4.	After work, I tend to need more time than in the past in order to relax and feel better *(E.R)*	*1*	*2*	*3*	*4*
5.	I can tolerate the pressure of my work very well *(E)*	*1*	*2*	*3*	*4*
6.	Lately, I tend to think less at work and do my job almost mechanically *(D.R)*	*1*	*2*	*3*	*4*
7.	I find my work to be a positive challenge *(D)*	*1*	*2*	*3*	*4*
8.	During my work, I often feel emotionally drained *(E.R.)*	*1*	*2*	*3*	*4*
9.	Over time, one can become disconnected from this type of work *(D.R)*	*1*	*2*	*3*	*4*
10.	After working, I have enough energy for my leisure activities *(E)*	*1*	*2*	*3*	*4*
11.	Sometimes I feel sickened by my work tasks *(D.R)*	*1*	*2*	*3*	*4*
12.	After my work, I usually feel worn out and weary *(E.R)*	*1*	*2*	*3*	*4*
13.	This is the only type of work that I can imagine myself doing *(D)*	*1*	*2*	*3*	*4*
14.	Usually, I can manage the amount of my work well *(E)*	*1*	*2*	*3*	*4*
15.	I feel more and more engaged in my work *(D)*	*1*	*2*	*3*	*4*
16.	When I work, I usually feel energized *(E)*	*1*	*2*	*3*	*4*

Note: Disengagement items are 1, 3(R), 6(R), 7, 9(R), 11(R), 13, 15. Exhaustion items are 2(R), 4(R), 5, 8(R), 10, 12(R), 14, 16. (R) means reversed item when the scores should be such that higher scores indicate more burnout.

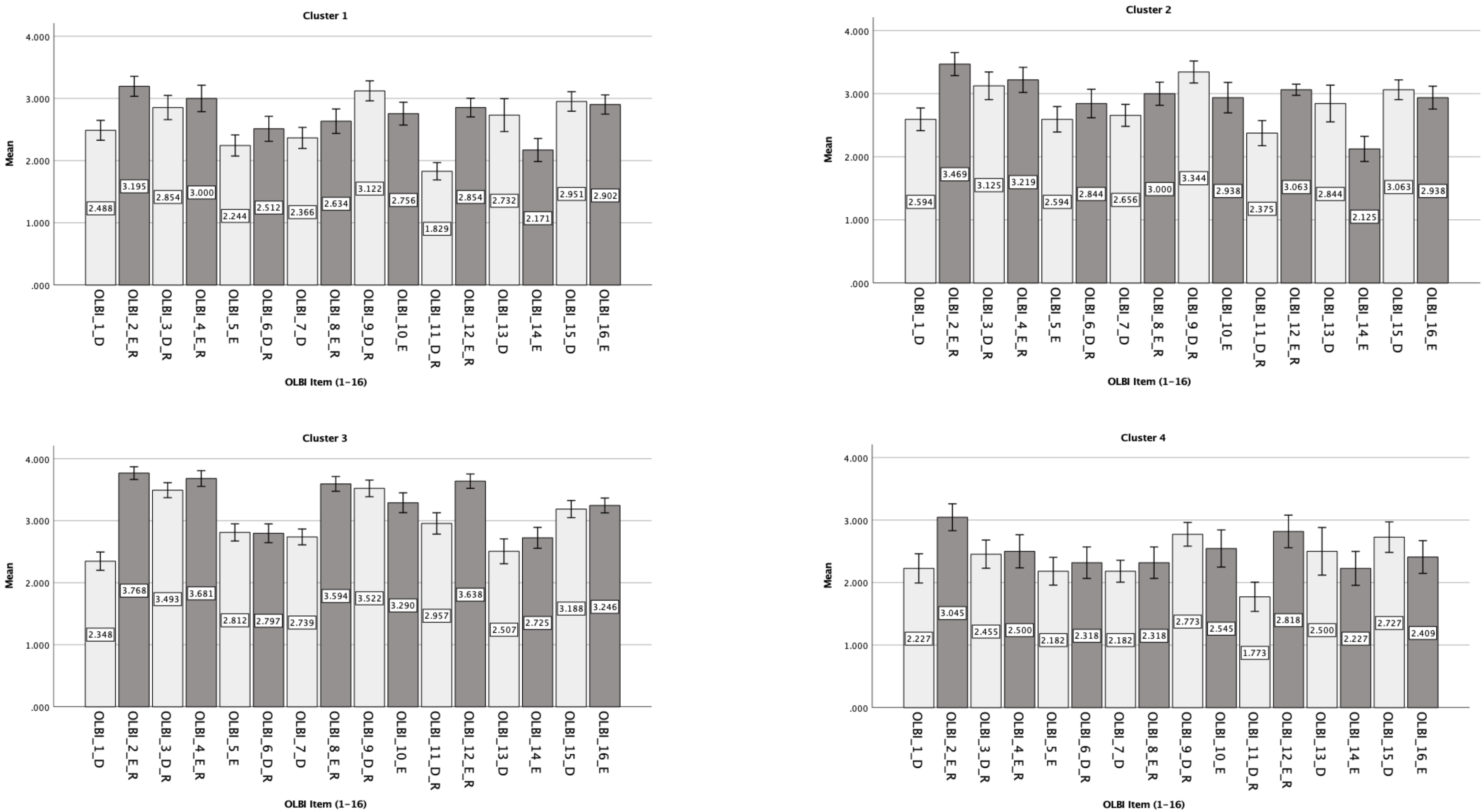
disengagement sub-total: ***exhaustion sub-total:*** ***full scale total:***

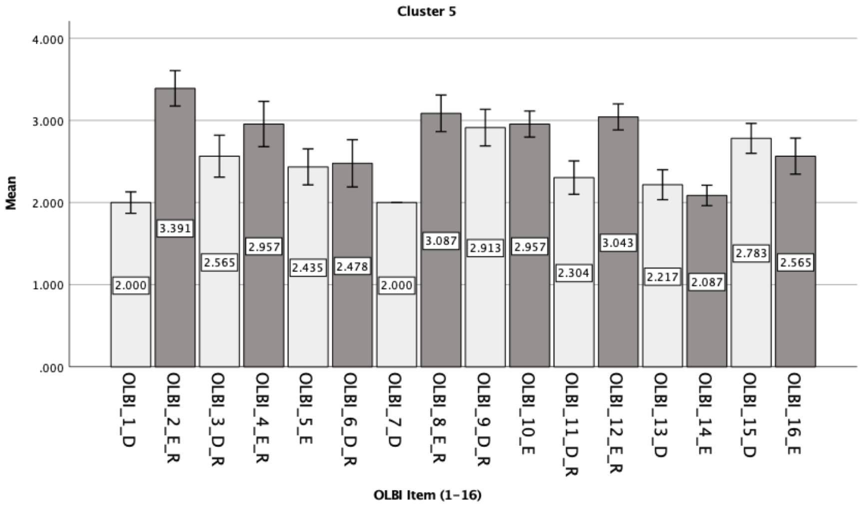
Delgadillo et al (2018) reported "Therapists are identified as having low, medium or high OLBI-D scores, based on scores above or below 1 standard deviation of the mean (M = 2.15, SD = 0.52; ≤1.62 = low, 1.63 to 2.67 = medium, ≥2.68 = high)."

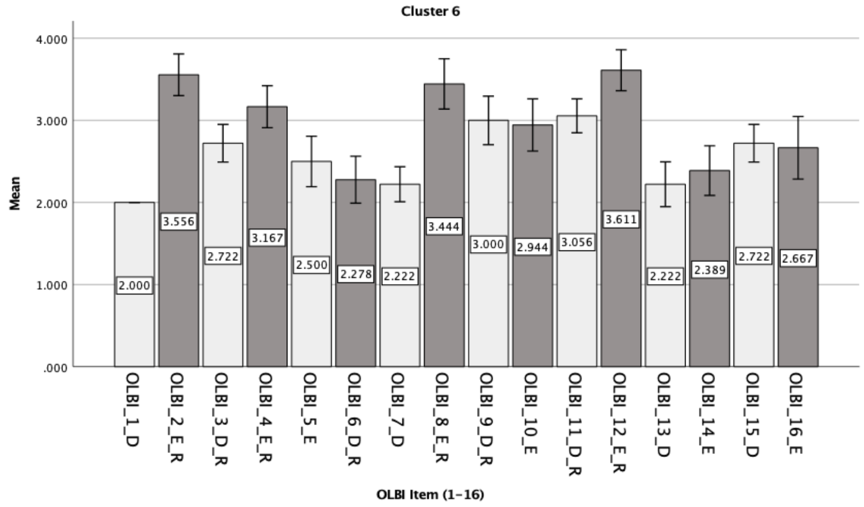
Appendix D

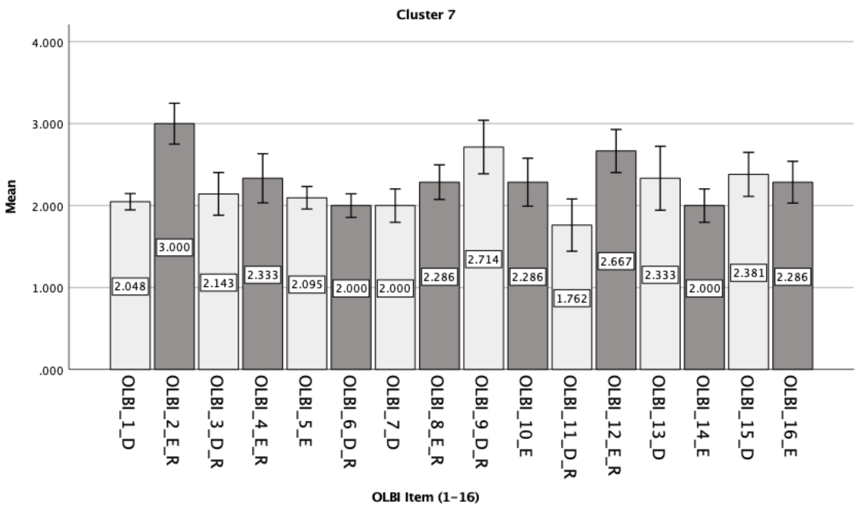
Figure 3

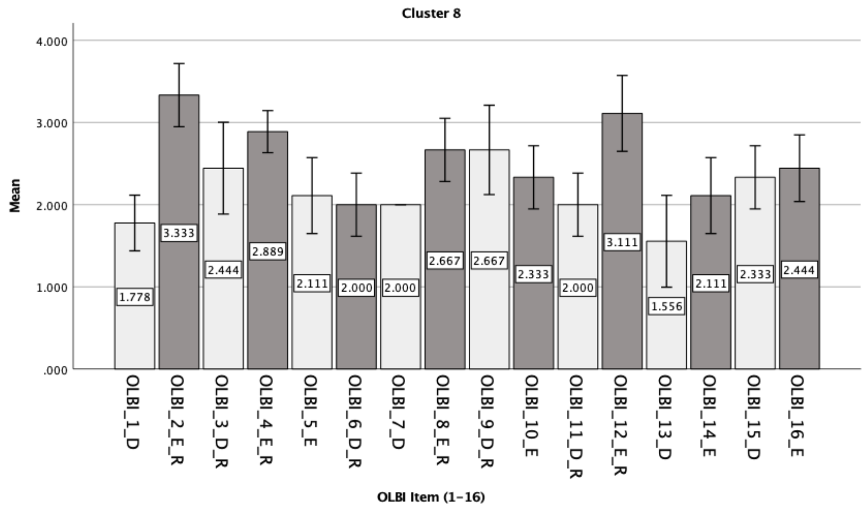
Mean Baseline Score of Each Item on the OLBI for Clusters 1-12 in the Test Sample (CPM 2)

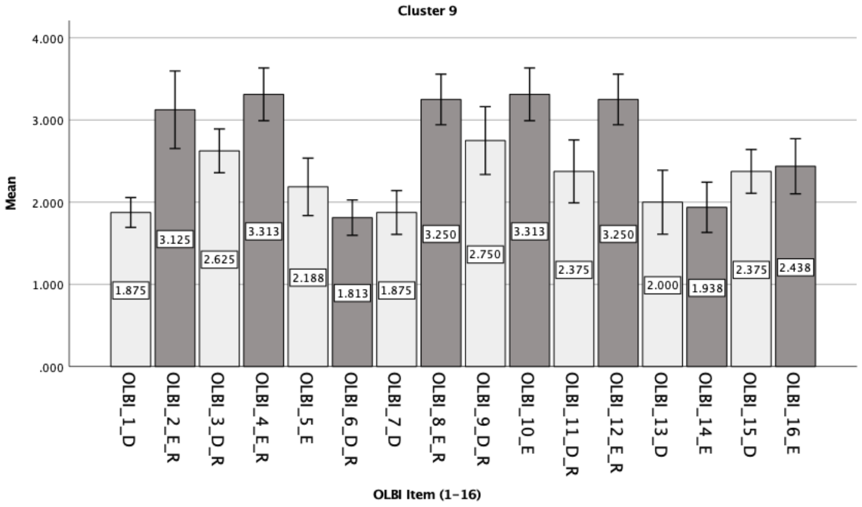


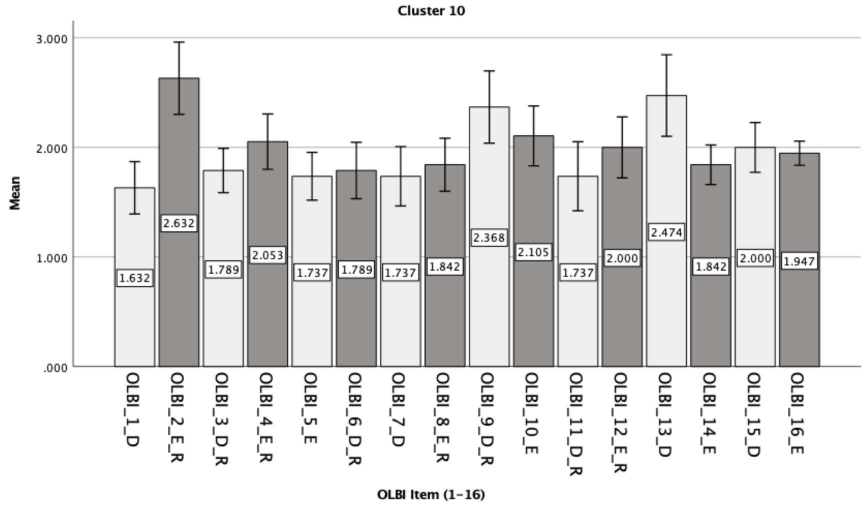

Cluster 5
Mean
OLBI Item (1–16)
2.000
3.391
2.565
2.957
2.435
2.478
2.000
3.087
2.913
2.957
2.304
3.043
2.217
2.087
2.783
2.565

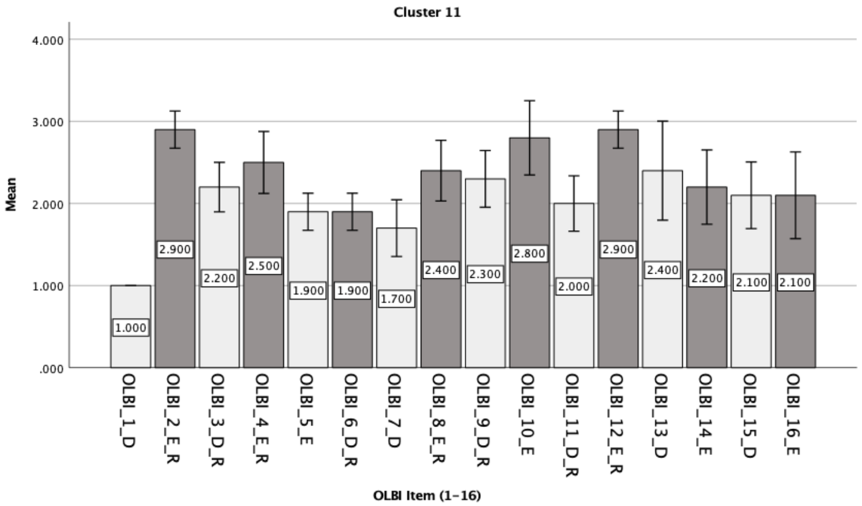

Cluster 6
Mean
OLBI Item (1–16)
2.000
3.556
2.722
3.167
2.500
2.278
2.222
3.444
3.000
2.944
3.056
3.611
2.222
2.389
2.722
2.667

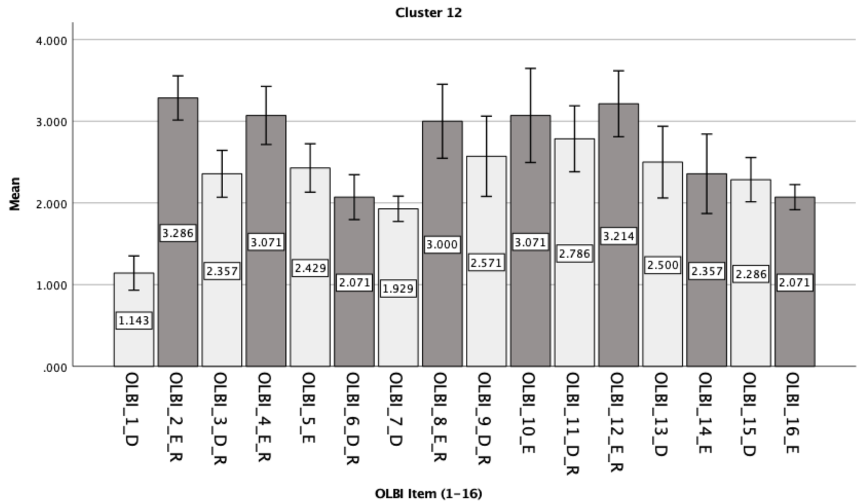

Cluster 7
Mean
OLBI Item (1–16)
2.048
3.000
2.143
2.333
2.095
2.000
2.000
2.286
2.714
2.286
1.762
2.667
2.333
2.000
2.381
2.286


Cluster 8
Mean
OLBI Item (1–16)
1.778
3.333
2.444
2.889
2.111
2.000
2.000
2.667
2.667
2.333
2.000
3.111
1.556
2.111
2.333
2.444


Cluster 9
Mean
OLBI Item (1–16)
OLBI_1_D 1.875
OLBI_2_E_R 3.125
OLBI_3_D_R 2.625
OLBI_4_E_R 3.313
OLBI_5_E 2.188
OLBI_6_D_R 1.813
OLBI_7_D 1.875
OLBI_8_E_R 3.250
OLBI_9_D_R 2.750
OLBI_10_E 3.313
OLBI_11_D_R 2.375
OLBI_12_E_R 3.250
OLBI_13_D 2.000
OLBI_14_E 1.938
OLBI_15_D 2.375
OLBI_16_E 2.438


Cluster 10
Mean
OLBI Item (1–16)
OLBI_1_D 1.632
OLBI_2_E_R 2.632
OLBI_3_D_R 1.789
OLBI_4_E_R 2.053
OLBI_5_E 1.737
OLBI_6_D_R 1.789
OLBI_7_D 1.737
OLBI_8_E_R 1.842
OLBI_9_D_R 2.368
OLBI_10_E 2.105
OLBI_11_D_R 1.737
OLBI_12_E_R 2.000
OLBI_13_D 2.474
OLBI_14_E 1.842
OLBI_15_D 2.000
OLBI_16_E 1.947


Cluster 11
Mean
OLBI Item (1–16)
OLBI_1_D 1.000
OLBI_2_E_R 2.900
OLBI_3_D_R 2.200
OLBI_4_E_R 2.500
OLBI_5_E 1.900
OLBI_6_D_R 1.900
OLBI_7_D 1.700
OLBI_8_E_R 2.400
OLBI_9_D_R 2.300
OLBI_10_E 2.800
OLBI_11_D_R 2.000
OLBI_12_E_R 2.900
OLBI_13_D 2.400
OLBI_14_E 2.200
OLBI_15_D 2.100
OLBI_16_E 2.100


Cluster 12
Mean
OLBI Item (1–16)
OLBI_1_D 1.143
OLBI_2_E_R 3.286
OLBI_3_D_R 2.357
OLBI_4_E_R 3.071
OLBI_5_E 2.429
OLBI_6_D_R 2.071
OLBI_7_D 1.929
OLBI_8_E_R 3.000
OLBI_9_D_R 2.571
OLBI_10_E 3.071
OLBI_11_D_R 2.786
OLBI_12_E_R 3.214
OLBI_13_D 2.500
OLBI_14_E 2.357
OLBI_15_D 2.286
OLBI_16_E 2.071